One share fits all? Regional variations in the extent of shadow economy in Europe*

Helmut Herwartz† Friedrich Schneider‡ Egle Tafenau§

February 8, 2010

Abstract

For the first time in this paper a multiple indicators multiple causes approach, amended to include spatial effects, is adopted to estimate the extent of the shadow economy in the European Union at the NUTS 2 regional level. It turns out that in the year 2004 the share of shadow economy was smallest in the regions of Netherlands, below 10%, while the Polish regions faced the largest share of shadow economy, around 30%. Our results are in general consistent with country level estimates from earlier studies. The variation of the extent of the shadow economy is in some countries considerable. Thus, policy measures against shadow activities should take the specific regional situation into account. Moreover, in implementing the regional policies of the European Union interactions with the shadow economy should be considered.

JEL-Classification: O17, C39, H26

Keywords: Shadow economy, European Union, regions, MIMIC modelling, spatial effects

---

*Earlier versions of the paper have been presented at the 2009 Annual Conference of the Estonian Economic Association, in the research seminar at the Institute of Statistics and Econometrics at the Christian-Albrechts-University of Kiel and the conference The Shadow Economy, Tax Evasion and Social Norms in Münster, Germany. We are grateful to the participants for their helpful remarks. Special thanks to Karsten Staehr for his thorough comments.

†Institute of Statistics and Econometrics, Christian-Albrechts-University of Kiel, D-24098 Kiel, Germany.

‡Department of Economics, Johannes Kepler University of Linz, A-4040 Linz-Auhof, Austria.

§Corresponding author. Address: Institute of Statistics and Econometrics, Christian-Albrechts-University of Kiel, D-24098 Kiel, Germany. E-mail: tafenau@bwl.uni-kiel.de.

1 Introduction

The shadow economy is by its nature difficult to measure: agents engaged in shadow activities try to avoid getting detected. Various techniques have been developed for estimating the size of shadow economy, since governments and economists are interested in knowing the true extent of economic activity. On the one hand, such knowledge is helpful for a better design of policies depending on the general economic situation (e.g. stabilisation policies). On the other hand, the shadow economy quota is a core input to estimate the amount of missed tax revenues and to determine a plausible extent of control for detecting unofficial economic activity.

Several problems arise when measuring shadow economy. Firstly, it is difficult to measure something that is hidden. Secondly, measurement of the shadow economy has been accused to be done without reliance on theory, which has given rise to the notion of 'guestimating' (Thomas 1999). Thirdly, even though selected authors provide an explicit definition of the shadow economy, it is often not clear if the constructed quotas correspond to this definition, especially if macroeconomic data are used (Williams & Windebank 1998).

The difficulties with measuring the shadow economy start from the term itself and its interpretation. There is no consensus of what the shadow economy exactly is and how it should be called (hidden economy, unofficial economy, informal economy, black work etc.; see Williams (2004) for an overview). Numerous interpretations are possible on what should be understood as the shadow economy, depending on the issue at hand. Most often the term 'shadow economy' refers to economic activities that are included in the calculation of gross domestic product (GDP) but are unregistered at governmental authorities (Chaudhuri et al. 2006). This comes from the presumption that the lead purpose of engagement in shadow activities is to avoid taxes and, accordingly, agents do not report their respective activities. Authorities accuse those involved in shadow activities for the loss in tax revenues, on account of which the government has a lowered ability to supply public services, or to disregard labour market, safety or quality regulations. On the other hand, shadow economy has also positive aspects. For instance, it works as a buffer for the labour force in times of economic downturn. Nevertheless, also at the level of the European Union (EU) the need to fight against undeclared work has been expressed (COM(2007) 628 final).

The size of shadow economy can be measured based on surveys or by applying indirect methods. In the latter case, the analysis is usually conducted by means of country level time series data, especially in case of multiple indicators multiple causes (MIMIC) modeling, which is also adopted in this work. In this approach the latent shadow economy is formalized as the outcome of its causes like tax rates and the degree of regulation or unemployment on the one hand. On the other hand, there are variables thought to indicate the size of the shadow economy, for instance, currency ratios or the rate of labour force participation. The distinction between causes and indicators is not clear cut, several variables could be thought of as both causing the shadow economy and responding to its size.

Regional variations of the shadow economy quota have been analysed only in a few cases though it could deliver useful input for designing policies for regional development and combating tax evasion. Mróz (2005) and Grabowski (2003) report results on the regional distribution of informal labour in Poland, based on surveys. Similarly, Schneider (2003) calculates the extent of shadow activities for the federal states of Germany. The MIMIC approach is applied in Chaudhuri et al. (2006) for Indian states.

In this work, for the first time, the size of the shadow economy is estimated for 238 EU regions at the NUTS 2 classification level.[1] Moreover, it is the first application of MIMIC modeling where spatial effects are taken into account. A further contribution is a new calibration procedure based on the sample mean and variation. This approach avoids the sensitivity of the estimates with respect to the chosen base observations. Finally, we consider the effect of shadow economy on the eligibility for finances from the EU regional policy budget under the Convergence objective.

[1]NUTS—The Nomenclature of Territorial Units for Statistics.

To preview some results, the empirical evidence reveals that in some countries the variation of the extent of shadow activities is considerable, such that regionally differentiated policies might be required to combat shadow economy. The calculated national average shadow economy quotas are in general consistent with previous estimates from country level analyses. As for the EU regional policy under the Convergence objective, it is argued that the effect to the shadow economy should be included in deciding the financing of proposed projects for economic development.

The remainder of the paper is organised as follows. Section 2 outlines shortly MIMIC estimation, followed by the description of the model, the data and the model estimation results in Section 3. Section 4 introduces the calibration technique, provides the estimates of the shadow economy quotas and a discussion of the regional variation of shadow activities. In Section 5 the implications for the EU regional policy are discussed. Section 6 provides a summary and some policy conclusions.

2 The MIMIC approach

The methods for estimating the extent of shadow economy range from relatively simple physical input methods like estimates based on electricity consumption to more complicated latent variable approaches. Survey data on tax evasion have also been used to determine the shadow economy quotas. Williams (2004) argues that surveys are the most reliable source of information on 'cash-in-hand work', finding that claims of using and supplying that kind of work give fairly the same results. Schneider & Enste (2000) show that using survey data leads to an underestimation of the size of the shadow economy, while estimates based on currency ratios or electricity consumption overestimate it (Schneider & Enste 2000, Giles 1999).

According to Giles (1999) and Schneider & Enste (2000), currently the MIMIC approach is considered to be the most reliable for estimating the extent of the shadow economy as it exploits simultaneously the informational content of both its causes and its indicators. The first application of MIMIC modelling to measure shadow economy was Frey & Weck-Hanneman (1984), but some aspects of the MIMIC model had been used earlier to determine the relative size of the shadow economy in different countries (Frey & Weck 1983).

The method has also earned criticism. Helberger & Knepel (1988) criticise the work of Frey & Weck-Hanneman (1984), showing that the results are highly sensitive with respect to variations in the sample. Moreover, they question whether the latent variable can be interpreted as shadow economy. Eilat & Zinnes (2000) argue that the MIMIC approach is more suitable to determine the 'potential for shadow activity' rather than to quantify its real extent. A critical discussion of applying the method in the context of shadow economy is provided by Breusch (2005). For a detailed literature overview on the measurement and size of the shadow economy the reader is referred to Schneider & Enste (2000).

MIMIC modelling is based on Zellner's (1970) approach of estimating regressions with unobserved independent variables, later generalized by Jöreskog & Goldberger (1975). It belongs to a group of models that consist of linear structural relationships allowing unobservable variables. In our context, the unobservable (latent) variable is the extent of the shadow economy.

The core of a MIMIC model consists of a structural equation and measurement equations. The structural equation relates the cause variables to the unobserved latent variable

$$\eta_r = \boldsymbol{\gamma}'\mathbf{x}_r + \zeta_r, \tag{1}$$

where η_r stands for the latent variable (shadow economy) in region r and the $(q \times 1)$ vector $\mathbf{x}_r = (x_{1r}, x_{2\,r}, \ldots, x_{qr})'$ collects the causes of shadow activities. The corresponding parameters are denoted by $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_q)'$ and ζ_r is an error term. All variables are measured as deviations from their means, thus, $\mathrm{E}[\mathbf{x}_r] = \mathbf{0}$ and $\mathrm{E}[\eta_r] = \mathrm{E}[\zeta_r] = 0$. Moreover, the error term ζ_r is assumed to be uncorrelated with the causes, i.e. $\mathrm{E}[\mathbf{x}_r\zeta_r] = \mathbf{0}$, its variance is $\mathrm{Var}[\zeta_r] = \psi$, and the covariance matrix of the cause variables is $\mathbf{E}[\mathbf{x}_r\mathbf{x}_r'] = \mathbf{\Phi}$.

The measurement equations relate indicators to the latent variable,

$$\mathbf{y}_r = \boldsymbol{\lambda}\eta_r + \boldsymbol{\varepsilon}_r, \tag{2}$$

where $\mathbf{y}_r = (y_{1r}, y_{2r}, \ldots, y_{pr})'$ is a $(p \times 1)$ vector of indicators of the latent variable and $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_p)'$ are the model parameters quantifying the impact of the shadow economy onto its indicators. The white noise error terms in the measurement equations are denoted as $\boldsymbol{\varepsilon}_r = (\boldsymbol{\varepsilon}_{1r}, \boldsymbol{\varepsilon}_{2r}, \ldots, \boldsymbol{\varepsilon}_{pr})'$. Again, $\mathrm{E}[\mathbf{y}_r] = \mathrm{E}[\boldsymbol{\varepsilon}_r] = \mathbf{0}$ and $\mathrm{E}[\mathbf{x}_r\boldsymbol{\varepsilon}_r'] = \mathbf{0}$, $\mathrm{E}[\boldsymbol{\varepsilon}_r\eta_r] = \mathbf{0}$, $\mathrm{E}[\boldsymbol{\varepsilon}_r\zeta_r] = \mathbf{0}$ by assumption. Finally, the covariance structure of the disturbances in the measurement equations (2) is given by $\mathrm{E}[\boldsymbol{\varepsilon}_r\boldsymbol{\varepsilon}_r'] = \boldsymbol{\Theta}_\varepsilon$.

In order to estimate the model, a theoretical covariance matrix of the observed variables is derived based on the reduced form of the model. In estimating the model parameters, those are found such that the difference between the theoretical covariance matrix and the empirical covariance matrix as calculated from the sample data is minimized. Details on the estimation can be found in Appendix A.

3 The model, data and parameter estimates

3.1 The model

Our model set-up relies on the approach followed in Dell'Anno et al. (2007). Their analysis is based on time series data for France, Spain and Greece. A similar model has been used in Dell'Anno (2007) for measuring shadow economy in Portugal. Full correspondence with these studies is not possible, however, due to unavailability of money demand at the regional level.

Two indicators of the shadow economy quota are employed, official GDP per capita and the rate of labour force participation. The former is expected to have a negative relation with the extent of the shadow economy and is the normalising variable, i.e. its coefficient is fixed to -1. The labour force participation rate is also likely negatively related to the extent of shadow economy since we expect that the more agents are engaged in official labour markets the less are concerned with shadow activities.

Six cause variables are distinguished. The first group of variables consists of direct and indirect taxes. In general, higher taxes are expected to encourage shadow activities. However, if instead of the imposed tax rates, effective tax rates are used—i.e. the share of collected tax revenue in the tax base—the data reflect also tax compliance. Therefore, depending on the data used, the relation of the tax variable to the shadow economy might be ambiguous. Second, the goodness of public services or the scope of public control (on tax evasion, sticking to regulations) is employed which should be negatively related to the extent of shadow activities. People are more willing to pay taxes if their contributions are used for good public services or the risk of being detected is high in case of participation in the shadow economy. The third group of causes, consisting of unemployment and self-employment rates, indicates labour market conditions. Both variables are expected to have a positive effect on the extent of shadow economy. Unemployed persons have more time to engage in shadow activities than people with full-time jobs. For the self-employed it is easier to hide their true incomes than it is for employed persons.

3.2 Data

The variables used in Dell'Anno et al. (2007) and our counterparts to those are documented in table 1. Moreover, the table documents the regional level at which the data are available. The data are retrieved from the Eurostat regional database, except for the tax wedges (drawn from Eurostat 2006) and the value added tax (VAT) rates (European Commission 2009). The year of analysis is 2004. All variables are used in the form of their relative difference from sample averages.

Considering the correspondence of the indicators to their theoretical counterparts, some comments are in order. Notably the imposed tax rates reflect neither the progressiveness of the tax system nor the differences in the allowed deductions. The effective

Table 1: Causes and indicators of shadow economy

Variables (Dell'Anno et al. 2007)	Measures	Regional level
	Indicators	
Real GDP per capita	GDP in purchasing power standards (PPS) per inhabitant	NUTS 2
Labour force participation ratio	Economic activity rate (15 to 64 years)	NUTS 2
Currency ratio	Not applicable in regional context	
	Causes	
Direct tax / GDP	- Paid taxes[a]	NUTS 2
	- Tax wedge[b]	NUTS 0
Indirect tax / GDP	Value added tax (VAT) rate	NUTS 0
Social security contributions / GDP	Included in the paid taxes and tax wedge measures[c]	
Public employment / Labour force	Employment by economic activity: NACE[d] sectors L to Q / Total	NUTS 2
Unemployment rate	Unemployment rate (15 years and over)	NUTS 2
Self-employment / Labour force	Number of self-employed / Labour force	NUTS 2

[a] Calculated using the data from households accounts: Paid taxes = (current taxes on income, wealth, etc. + social contributions) / balance of primary income. The balance of primary income is the income earned by the households (profits from self-employment, wages, property income) minus property costs (rents, interests, etc.), without subtracting taxes and adding transfers. The tax measures come from the secondary distribution of income account of households. It relies on the balance of primary income, subtracts the taxes paid and adds transfers received by the households, in order to derive the disposable income. Thus, the paid taxes measure shows the share of paid direct taxes in households' gross income.

[b] Difference between the labour cost for an employer and the net wage his employee takes home. Calculated for a single worker without children at 2/3 of average earnings.

[c] Dell'Anno et al. (2007) compare models with varying aggregation levels of the tax variables. In the most general specification, they differentiate between direct taxes, indirect taxes and social contributions. In the most aggregated version a measure on general tax burden is used.

[d] Classification of Economic Activities in the European Community.

tax rate (paid taxes) compensates for this issue to some extent. Moreover, the share of public employment has been stated to measure the goodness of public services or the scope of public control. In fact, this measure can also indicate ineffectiveness of the public sector. In this case its expected relationship with the shadow economy is positive since people are not motivated to pay taxes for financing an overly large public sector.

The initial sample includes all NUTS 2 regions of the 27 EU member states. Mainly due to missing data, some of these have been removed. Firstly, the oversea territories are omitted. Secondly, Bulgaria, Cyprus, Luxembourg, Malta and Romania have been dropped due to missing data for some of the variables. For the same reason two regions of the United Kingdom, North Eastern Scotland and Highlands and Islands, are excluded. Finally, Denmark and Slovenia are not represented with their NUTS 2 regions, but with the country as a whole since only country level data are available. After these adjustments, 238 regions remain in the sample. In case of some still missing observations the data for corresponding NUTS 1 regions are used, if available, or national data if also the data for NUTS 1 regions are missing. The year of the analysis is 2004, chosen as the year with least missing data.

3.3 The estimated model

For estimating the model, we have tried to use ML, GLS and ULS algorithms outlined in Appendix A. The results from the GLS estimation are documented in table 2.[2] The full model is presented as model 1. In models 2–7, there is always one of the cause variables omitted, to assess the robustness of the results. As regional data are known to be spatially correlated, in models 8, 9 and 10 spatial effects are taken into account.[3] In models 8 and 9 the indicator variables are spatially adjusted: in model 8 with the same

[2]ML estimation failed due to singularity problems. ULS coupled with bootstrap inference delivers results similar to those of GLS. The only qualitative difference is the insignificance of tax wedge, share of public employment and self-employment rate in the ULS estimation. ULS estimation results are available from the authors on request.

[3]The procedure is described in Appendix B. In the context of shadow economy, spatial dependencies might arise if the decision to participate in shadow activities can be influenced by employment possibilities in neighbouring regions.

spatial effects coefficient and in model 9 with distinct ones. In model 10, it is assumed that the cause variables have to be adjusted for spatial effects, using an identical spatial effects coefficient. The estimation results are mostly discussed with regard to a 5% significance level.

All of the estimated models have a very good fit, as indicated by the high values of the adjusted goodness-of-fit index. The χ^2-test of overall model fit conforms with the hypothesis of equal implied and sample covariance matrix with 5% significance level.

The parameter estimates have their expected signs and are significant. However, in discussing the model set up and data the effect of the measure of paid taxes and public employment were marked to be ambiguous. The share of paid direct taxes in the income of the households can indicate both tax burden and tax compliance. If the tax compliance argument prevails, this measure should have a negative impact on the extent of shadow economy, as indicated by our estimates. This effect is somewhat magnified by including the imposed tax rate in the model (compare the results from models 1 and 3).[4] In earlier studies of the country level shadow economy, the effective tax rate, calculated as the share of collected tax revenues in GDP, has been found to be positively related to the extent of shadow economy (Dell'Anno et al. 2007).

The tax wedge and the VAT rate have a positive effect on the shadow economy quota. Especially in case of the VAT the effect is very stable across the alternative model specifications. In spite of the ambiguity of the interpretation of the public employment variable, its effect is positive and stable across the different model specification. Thus, it can be argued to show public sector inefficiency. The only exception is model 9, in which its effect is insignificant. The estimation results of Dell'Anno et al. (2007) are contradictory, depending on the country of analysis.

Both labour market variables, the unemployment rate and the self-employment rate, have a significantly positive relation with the extent of shadow economy. However, their

[4] If in addition a measure on the tax paying moral would be included, the coefficient would probably be smaller. Such an indicator is, however, not available for the whole sample.

Table 2: MIMIC parameter estimates

Variables	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
Causes										
Paid taxes	−0.598*		−0.501*	−0.536*	−0.518*	−0.722*	−0.700*	−0.593*	−0.522*	−0.597*
	(−5.89)		(−5.01)	(−5.06)	(−5.39)	(−6.51)	(−6.56)	(−5.59)	(−5.01)	(−5.92)
Tax wedge	0.363*	0.089		0.463*	0.304*	0.548*	0.233*	0.351*	0.247*	0.321*
	(4.05)	(0.92)		(4.26)	(3.46)	(5.79)	(2.37)	(4.22)	(2.81)	(3.52)
VAT	0.548*	0.526*	0.640*		0.577*	0.492*	0.643*	0.610*	0.615*	0.548*
	(5.64)	(5.23)	(6.33)		(6.05)	(5.25)	(6.44)	(6.22)	(6.29)	(5.71)
Share of public employment	0.193*	−0.072	0.113	0.205*		0.267*	0.045	0.163*	0.078	0.171*
	(2.73)	(−0.94)	(1.45)	(2.50)		(4.00)	(0.58)	(2.28)	(1.02)	(2.27)
Unemployment rate	0.128*	0.227*	0.190*	0.124*	0.151*		0.164*	0.138*	0.148*	0.133*
	(4.13)	(6.03)	(6.05)	(3.37)	(4.41)		(4.45)	(4.74)	(4.39)	(4.26)
Self-employment rate	0.154*	0.188*	0.127*	0.176*	0.129*	0.165*		0.171*	0.150*	0.144*
	(4.53)	(4.91)	(3.81)	(4.57)	(3.92)	(4.84)		(4.70)	(4.89)	(4.76)
Indicators										
GDP (PPS) per capita	−1.0	−1.0	−1.0	−1.0	−1.0	−1.0	−1.0	−1.0	−1.0	−1.0
Labour force participation	−0.284*	−0.263*	−0.255*	−0.294*	−0.267*	−0.325*	−0.249*	−0.33*	−0.459*	−0.271*
	(−6.77)	(−7.10)	(−7.37)	(−6.33)	(−6.56)	(−6.69)	(−6.69)	(−6.96)	(−6.80)	(−7.41)
Statistics										
Degrees of freedom	25	18	18	18	18	18	18	25	25	25
χ^2	33.07	30.88	14.05	32.69	18.91	19.48	18.25	32.07	22.73	23.82
RMSR	0.004	0.004	0.003	0.005	0.003	0.003	0.002	0.004	0.004	0.003
AGFI	0.950	0.942	0.974	0.939	0.965	0.963	0.966	0.951	0.965	0.964
R^2_η	0.539	0.693	0.362	0.688	1.031	−0.036	0.548	0.864	0.983	0.973
Spearman's rank correlation of η	1.000	0.907	0.956	0.906	0.986	0.929	0.946	0.998	0.991	0.992

Source: Own calculations. GLS estimation results. Models 8–10 adjust for spatial effects. t-statistics in parentheses. * denotes $|t\text{-statistic}| > 1.96$. GDP per capita is the normalising variable with its coefficient fixed to −1. RMSR—root mean squared residual, AGFI—adjusted goodness-of-fit index, $R^2_\eta = 1 - \phi/\operatorname{Var}[\eta]$ —coefficient of determination for the latent variable.

effect is smaller than the effect of the taxes. Such a direct comparison is possible as the size of the coefficients is relative to the normalizing variable GDP per capita.

From the indicator variables the labour force participation rate is negatively related to the shadow economy. As mentioned, GDP per capita is chosen as the normalising variable. The correspondence of the signs of other parameter estimates to their expected direction indicates that normalisation to a negative value is justified.

Omitting one of the cause variables from the model specification has only minor effects on the parameter estimates as can be seen from models 2–7. In most of the cases there are changes neither in the signs nor in the significance of the estimates. The exceptional cases are models 2, 3 and 7. In models 3 and 7, the only difference is the insignificance of the share of public employment. In model 2 where the regionally available share of direct taxes in households incomes has been omitted, in addition the effect of the tax wedge is insignificant.

In models 8–10 the data have been adjusted for spatial effects. It is widely believed that in working with regional data, spatial effects from neighbouring regions should be considered (e.g. in analysing regional convergence in the EU Le Gallo & Dall'erba 2006). The decision whether to participate in the official or in the shadow economy in some region depends also on the economic possibilities in the neighbouring regions. E.g., it can be expected that shadow economy in a region is smaller if the neighbouring regions have high GDP per capita, such that it is possible to find an official job there if one is prepared to commute. This consideration would mean that "effective" GDP per capita indicating the extent of the shadow economy should be adjusted upwards (meaning a negative spatial effects coefficient ρ; see Appendix B). With respect to the labour force participation, the effect can be twofold. On the one hand, it is probably more difficult to get an official job in a neighbouring region with a high labour force participation rate due to sufficient domestic labour supply. On the other hand, a high rate of labour market participation can indicate good labour market possibilities. Thus, the labour force participation rate should be adjusted downwards or upwards. Similar considerations could be made for the cause variables.

In models 8 and 10 an identical spatial effects coefficient is employed for all indicator and cause variables, respectively. The estimation procedure delivered as the most appropriate spatial effects coefficient in case of the indicators -0.06 and in case of causes 0.10. When the spatial effects coefficient is allowed to differ in case of the two indicators, GDP per capita of the neighbouring regions turns out to be irrelevant ($\rho = 0$). The argument of good labour market conditions dominates in case of labour force pariticipation as shown by the negative spatial effects coefficient $\rho = -0.18$.

The inclusion of spatial effects reduces the estimation uncertainty of the latent variable as indicated by the high values of the coefficient of determination for the latent variable R^2_η in the spatial models. Thus, combining the MIMIC approach with spatial effects enables to obtain exacter latent variable estimates. However, the parameter estimates are not strongly influenced by including spatial effects. As the only exception, in model 9 public employment loses its significance. Also the resulting ordering of the regions according to the extent of shadow economy (discussed in the next section) is highly correlated in the spatially weighted and in spatially unweighted models, as indicated by the Spearman's rank correlation in table 2.

4 Shadow economy quotas and tax gaps

4.1 Calibrating shadow economy quotas

4.1.1 Calibration techniques used in the literature

Due to the normalisation, the MIMIC method obtains only a preliminary index for the extent of the shadow economy. To derive explicitly shadow economy quotas several approaches have been employed in the literature. Whatever the specific method, one or two base values are needed for the calibration. These can be taken from previous research as done, for instance, in Chaudhuri et al. (2006), Buehn & Schneider (2008), Frey & Weck-Hanneman (1984). Some other authors calculate an alternative set of estimates for the shadow economy quota themselves (Giles 1997).

The MIMIC modelling of the shadow economy has until now been used in time series (Giles 1997, Dell'Anno 2007, Dell'Anno et al. 2007, Buehn & Schneider 2008) or in a pooled cross section time series context (Frey & Weck-Hanneman 1984, Schneider & Bajada 2005, Schneider 2007). If the latent variable is assumed to be directly related to the percentage share of the shadow economy in the GDP, one needs two base values: one for fixing the overall extent of the shadow economy and the other one for the step size (Frey & Weck-Hanneman 1984). If the latent variable can be interpreted as the growth rate of the shadow economy (this is the case if it is directly related to the GDP growth), only one base value is needed. Then, the remaining shadow economy quotas are found by integration (Schneider & Enste 2000, Dell'Anno et al. 2007). These two approaches are probably the simplest, but not the only ones. The way of calibrating the MIMIC index has caused controversies and there is still no unanimously accepted procedure (Breusch 2005, Dell'Anno & Schneider 2006).

In addition to choosing an appropriate calibration procedure, problems arise due to estimation errors inherent in the preliminary index obtained from equation (1). However, this issue has rarely earned a remark, with the exception of Giles & Tedds (2002). Finally, the reliability of the external or self-derived base value employed in the calibration should be carefully considered. Some authors (Dell'Anno & Schneider 2003) have used the average of all available estimates for a year and a country, hoping to avoid a bias in this way.

The calibration techniques relying on base values for one or two arbitrary individual observations are sensitive to measurement errors in the base values and to the chosen observations. Therefore, an alternative approach for choosing the base values is proposed here. It takes the average extent of the shadow economy and its variation in the sample as the points of reference. Even though the sample mean and the variation are subjected to measurement errors, the results are not as sensitive to a modification of initial parameters as in case of calibrating with item specific a priory quotas.

4.1.2 Calibration technique based on sample moments

When all variables in the model are measured in the form of relative differences from the sample average, it is reasonable to assume that this also holds for the latent variable. The preliminary latent variable index $\hat{\eta}_r$ calculated from equation (1) can be either assumed to be proportional to it or to show the relative difference itself. In the first case,

$$\hat{\eta}_r = a\frac{SE_r - \overline{SE}}{\overline{SE}}, \tag{3}$$

where SE_r is the region's true percentage share of shadow economy in the officially measured GDP, $\overline{SE}$ is the corresponding sample mean and a denotes a proportionality factor. Rewriting (3) for SE_r gives

$$SE_r = \overline{SE}\left(1 + \frac{\hat{\eta}_r}{a}\right). \tag{4}$$

Thus, in order to transform the preliminary index to the shares of shadow economy in GDP, the average extent of shadow economy and the proportionality factor a need to be known. We assume that the sample average and standard deviation of the shadow economy quota are given. Then, based on equation (4), the variance of the extent of shadow economy is

$$\mathrm{Var}[SE] = \frac{\overline{SE}^2\,\mathrm{Var}[\hat{\eta}]}{a^2}. \tag{5}$$

Solving (5) for the proportionality factor a gives

$$a = \overline{SE}\sqrt{\frac{\mathrm{Var}[\hat{\eta}]}{\mathrm{Var}[SE]}}. \tag{6}$$

The second option is to assume that the preliminary index reflects the relative difference from the mean of the latent variable, i.e. the proportionality factor is unity, $a = 1$.

In that case, the shadow economy quotas can be calculated as

$$SE_r = \overline{SE}\,(1 + \hat{\eta}_r)\,. \tag{7}$$

In this case only the sample average is needed for calculating the shadow economy quota. This approach is comparable to the calibration method in Schneider & Enste (2000) and Dell'Anno et al. (2007).

4.2 Shadow economy estimates

The obtained parameter estimates can be used to calculate a preliminary index for the shadow economy by means of equation (1). The estimate for the variance of the error term in this equation is around $\hat{\phi} = 0.03$ in models without adjusting for spatial effects and between 0.001 and 0.006 in models with spatially adjusted variables. The preliminary index of the latent variable ranges approximately from -0.65 to 0.65, with the range depending on the model specification. Thus, the possible error in the shadow economy estimates is in fact large but it is smaller in case of the spatial models. However, as discussed below, the order of the regions and the shadow economy quotas appear plausible (see Appendix C) and in line with related country level studies. For example, the Western European regions are characterised by smaller shadow economy quotas than the Eastern European regions.

In calibrating the shadow economy quota from the preliminary index, 17.2% and 5.4% are used as the average and the standard deviation of the shadow economy quota, respectively. These estimates originate from the national estimates for the EU countries from Schneider (2007) (documented here also in table 3) as weighted averages. In deriving the weighted average and standard deviation, it has been assumed that each region within a country has the same shadow economy quota as the country itself.[5]

[5]This assumption clearly contradicts the purpose of the paper. However, the standard deviation of the size of the shadow economy across all European regions is probably not strongly influenced by this assumption as within-country variation can in some countries be larger but in some other countries smaller than the European average.

Then, regional GDPs have been used as weights for calculating the moments.

In table 3 the weighted average rates of shadow economy across the regions of each country are presented for the spatially uncorrected full model (model 1) and the spatial models (8–10), derived with the calibration procedure using two parameters. In case of the spatially uncorrected model, besides the GLS results also the ULS estimates are presented, as well as estimates based on the second calibration procedure that relies only on the sample mean. For comparison the results of Schneider (2007) as the mean for 2003/2004 and 2004/2005 are provided.

Comparing the ULS with GLS results reveals that the obtained shadow economy quotas are not sensitive with respect to the choice of the estimation method. The only country with a remarkable difference is Ireland. Also the two calibration procedures deliver fairly similar results, with the two parameter method obtaining a slightly higher variance of the shadow economy quota across the countries (and regions) than the procedure that relies only on the mean. Our standard deviation estimate probably overestimates the true standard deviation in the sample somewhat, due to the extremely high shadow economy shares obtained for the Baltic countries in Schneider (2007). The results from the spatial models are similar to those of the model without a spatial adjustment.

Juxtaposing the shadow economy estimates to those of Schneider (2007) reveals a high similarity of the results. Especially salient in this respect are the estimates for Finland, Ireland, Germany, Poland, Sweden and United Kingdom (based on the GLS estimation of model 1). The largest differences can be observed for Denmark, Estonia, Greece, Latvia, Lithuania and Slovenia (difference more than 5 percentage points). Comparing the results to some alternative estimates in the literature (see table 4), it can be concluded that in some cases our estimates probably underestimate the shadow economy quota (Slovenia, Greece), while in some other cases the results in Schneider (2007) are likely to overestimate the extent of shadow economy (Latvia, Lithuania, Slovenia).

Table 3: The national averages of the shadow economy in 2004, in % of GDP[a]

Country	M1	M1[b]	M1[c]	M8[d]	M9[e]	M10[f]	Schneider (2007)[g]
Netherlands	9.2	9.7	12.1	9.5	9.7	9.3	11.6
Denmark	10.8	12.5	13.1	11.4	12.0	11.1	16.4
United Kingdom	11.8	13.4	13.8	11.9	12.1	11.7	11.0
Austria	13.9	14.3	15.1	14.0	14.2	14.1	9.7
Ireland	14.3	19.0	15.4	14.8	15.9	15.1	14.5
Germany	15.1	14.5	15.9	14.9	14.6	15.2	15.7
Sweden	15.8	14.2	16.3	16.0	15.9	15.7	16.8
Finland	15.9	16.9	16.4	16.1	16.4	16.1	16.1
Slovenia	16.2	17.7	16.6	16.3	16.6	16.5	27.8
France	16.2	15.9	16.6	16.0	15.8	16.3	13.5
Estonia	16.3	18.7	16.6	16.2	16.6	16.7	38.7
Spain	16.5	18.7	16.7	16.5	16.9	16.9	20.9
Czech Republic	16.7	17.3	16.9	16.9	17.1	16.2	18.8
Belgium	17.4	15.1	17.3	17.3	16.9	15.9	20.0
Portugal	18.8	19.4	18.2	19.0	19.3	19.0	20.8
Italy	20.4	18.7	19.3	20.6	20.5	20.4	24.0
Greece	20.6	20.3	19.4	20.8	20.9	20.7	26.9
Latvia	21.0	21.6	19.6	20.6	20.4	21.2	39.9
Slovak Republic	22.0	24.9	20.3	21.9	22.3	22.3	18.7
Hungary	22.3	21.2	20.5	22.4	22.4	22.3	24.8
Lithuania	22.9	22.8	20.9	22.6	22.4	23.1	30.8
Poland	28.8	30.3	24.7	28.9	29.2	28.9	27.8

Source: Own calculations.

[a] Based on GLS estimates, if not indicated otherwise. Weighted average, the regional GDP shares have been used as the weights. The calibration procedure using average = 17.2 and standard deviation = 5.4 has been applied if not indicated otherwise. The countries are ordered according to the shadow economy quota obtained from model 1 (GLS, two parameter calibration procedure).

[b] Based on ULS estimates.

[c] The calibration procedure using only average = 17.2.

[d] Indicators spatially adjusted, with an identical spatial effects coefficient.

[e] Indicators spatially adjusted, each with its own spatial effects coefficient.

[f] Cause variables spatially adjusted.

[g] Average over Schneider (2007)'s estimates for 2003/2004 and 2004/2005.

Table 4: Some former estimates of the shadow economy quota, % of GDP

Country	Estimate	Year	Method	Source
Greece	26–27	2002	MIMIC	Dell'Anno et al. (2007)
Italy	16.1–18.1	2004		ISTAT (2008)
Latvia	20.8	2003	Expenditures	Bernotaitė & Piskunova (2005)
Latvia	24–25	2003	Money demand	Brēķis (2007)
Lithuania	20.8	2004	Survey of market participants	LFMI (2005)
Slovenia	18.8–23.0	2004	Employment discrepancy	Nastav & Bojnec (2007)

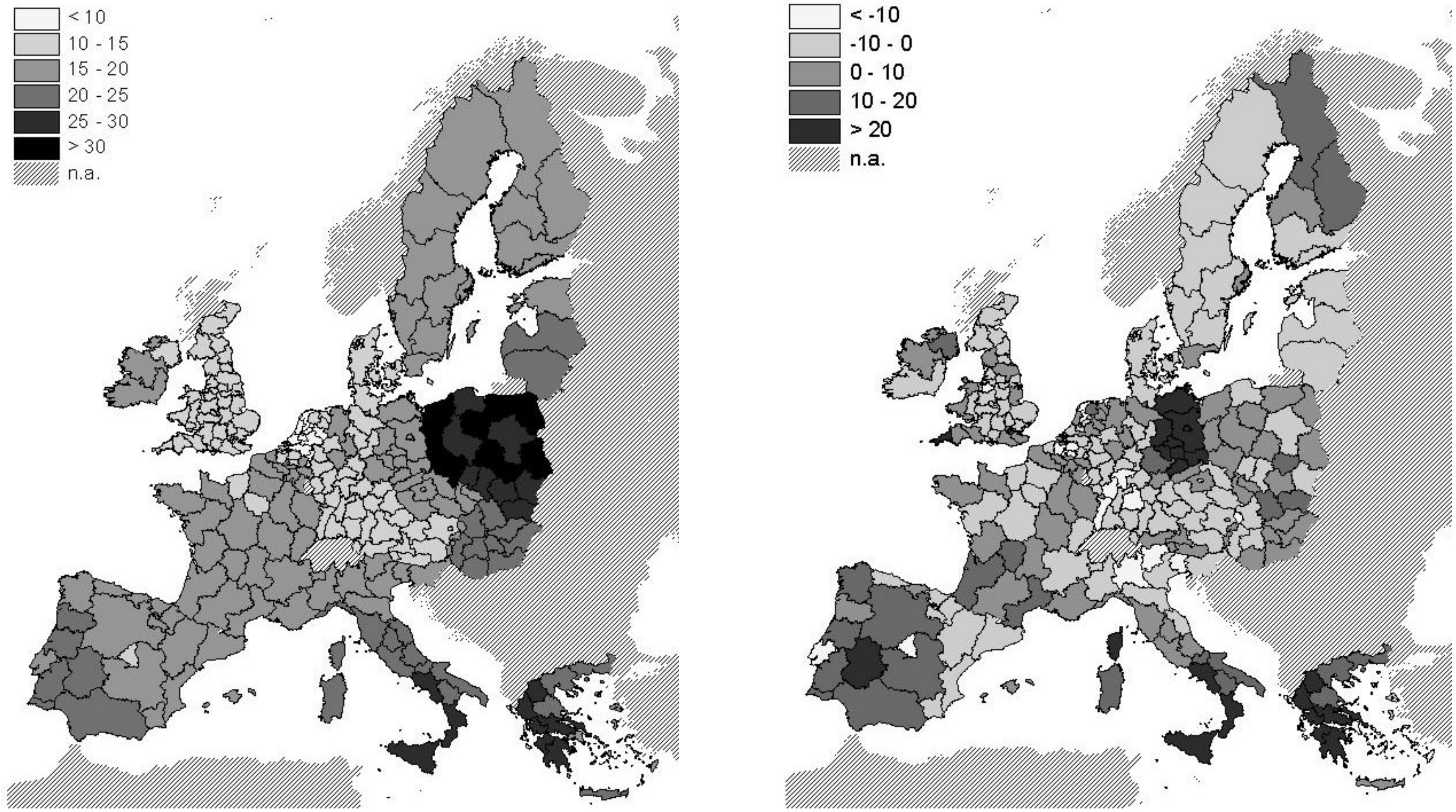
The regional estimates of the shadow economy quotas are given in Appendix C and presented in figure 1.[6] While panel (a) of the figure reveals that the share of shadow economy is highest in Eastern and Southern Europe, panel (b) indicates that the extent of shadow economy varies strongly also in several Western European countries.

The shadow economy quota is negatively correlated to the wealth of the region, in general relatively rich regions have a smaller shadow economy than relatively poor ones. This result holds both for the whole sample and within countries. For example, the regions in East Germany have higher shadow economy quotas than in West Germany.[7] However, there are some exceptions. First, according to the GLS estimation (model 1 and 10) the Greek region Ionia Nisia is estimated to have the smallest shadow economy in our sample, 5.9–6.3%.[8] In this region, the paid taxes variable takes the highest value in the sample, being also more than twice as high as in the rest of Greece. Therefore, it can be considered to be an outlier and our shadow economy estimate for this region is not reliable. According to the ULS estimation results as well as in case of spatially weighted indicator variables (model 9) the shadow economy quota in Ionia Nisia is still very low (10.2 and 9.5%, respectively), but the top position is achieved by Dutch regions with the quotas below 9%. Second, the prosperous regions Brussels and Inner London with

[6] We are grateful to Lorena Gola for preparing the maps.

[7] This result contradicts that of Schneider (2003) for 1995 and 1999: using survey data he concluded that the shadow economy quota was higher in the western part of Germany than in the eastern part.

[8] In fact, the smallest shadow economy is estimated for the Finnish region Åland, 1.8%. Considering this extremely low value, resulting from a zero VAT rate, the result is obviously unreliable.



(a) (b)

Figure 1: Shadow economy (a, % of GDP) and its deviation from the country average (b, %) in the EU NUTS 2 regions. n.a.—not available.

shadow economy quota around 20% and 14%, respectively, are regions with the highest or one of the highest shadow economy quota in Belgium and the United Kingdom. Behind this result are high unemployment rates. In fact, Brussels and Inner London have the highest unemployment rate within the respective countries.

In table 5 the coefficients of variation of the shadow economy quota are documented. The coefficient of variation shows how much the shadow economy quotas vary in a country across its regions and is calculated as the ratio of weighted standard deviation to the weighted average of the shadow economy quotas, using the regional GDP shares as the weights. The results reveal that there is considerable variation in the shadow economy quotas within some countries, the most outstanding in this respect are Belgium, Germany, Spain, Finland, Greece, Italy and the Slovak Republic. The governments of these countries should consider seriously whether 'one-size-fits-all' policies are reasonable in taking measures against shadow economy or if they should be regionally diversified. Relatively low variation in the shadow economy quotas are observed in Austria, the Czech Republic, Hungary and Sweden. Once again, these results are stable across different models.

5 Regional policy in the EU

One third of the budget of the EU (about 44 billion euro per year during the period 2007–2013) is spent on regional development, financed through the Cohesion Fund and Structural Funds. The respective finances are distributed under the objectives of Convergence, Regional competitiveness and employment, and European territorial cooperation. Most of these financial resources (81.54% during the period 2007–2013) are spent with regard to the Convergence objective (Council Regulation (EC) No 1083/2006, Article 19). In determining the eligibility of a region for finances under the Convergence objective its GDP per capita, adjusted for the purchasing power, plays a crucial role: eligible are regions whose GDP per capita (PPS) is below 75% of the EU average (Coun-

Table 5: The coefficients of variation of the shadow economy in 2004[a]

Country	M1	M1[b]	M1[c]	M9[d]	M10[e]	Number of regions
Austria	0.04	0.04	0.02	0.03	0.03	9
Belgium	0.15	0.18	0.10	0.14	0.18	11
Czech Republic	0.05	0.12	0.03	0.04	0.06	8
Germany	0.12	0.16	0.07	0.12	0.11	39
Denmark	0	0	0	0	0	1
Estonia	0	0	0	0	0	1
Spain	0.15	0.14	0.09	0.15	0.14	16
Finland	0.15	0.12	0.09	0.14	0.14	5
France	0.09	0.05	0.05	0.07	0.08	22
Greece	0.18	0.15	0.13	0.18	0.18	13
Hungary	0.05	0.04	0.03	0.04	0.04	7
Ireland	0.08	0.03	0.05	0.07	0.07	2
Italy	0.15	0.15	0.10	0.13	0.15	21
Lithuania	0	0	0	0	0	1
Latvia	0	0	0	0	0	1
Netherlands	0.06	0.05	0.03	0.05	0.08	12
Poland	0.07	0.08	0.05	0.07	0.07	16
Portugal	0.11	0.08	0.07	0.11	0.11	5
Sweden	0.03	0.04	0.02	0.03	0.03	8
Slovenia	0	0	0	0	0	1
Slovak Republic	0.13	0.19	0.09	0.14	0.13	4
United Kingdom	0.10	0.07	0.05	0.09	0.10	35

Source: Own calculations.

[a] Based on GLS estimates, if not indicated otherwise. Based on the weighted average and variance, the regional GDP shares have been used as the weights. The calibration procedure using average = 17.2 and standard deviation = 5.4 has been applied if not indicated otherwise.

[b] Based on ULS estimates.

[c] The calibration procedure using only average = 17.2.

[d] Indicators spatially adjusted, each with its own spatial effects coefficient.

[e] Cause variables spatially adjusted.

cil Regulation (EC) No 1083/2006, Article 5(1)).[9] Under the Regional competitiveness and employment objective, the eligible regions are those that do not qualify under the Convergence objective.

For a just classification of regions as eligible and non-eligible under the Convergence objective, it is important to assure that the official GDP figures cover the total productive activity, whether from the formal or shadow sector. Indeed, the EU requires its member countries to include all productive activities in the official GDP figures (ESA 1995).[10] In some sectors of the economy this can be achieved automatically due to the applied calculation methods of the volume of production or incomes. For example, in the United Nations' report 'Non-observed economy in national accounts' (United Nations 2008) agriculture and rents are mentioned for Germany as such sectors. However, the shares of hidden and observed economic activities remain unknown.

Statistical offices exert also direct adjustments of GDP in order to comprise unobserved economic activities. Those include illegal activities, shadow activities that are not captured by the implicit methods (deliberate non- or misreporting) and activities that are not required to be reported. The latter activities belong otherwise to the formal economy, as they pay their due taxes and comply to relevant regulations. According to United Nations (2008), such adjustments cover different aspects of unobserved economy across countries, being thus not fully internationally comparable. In some countries adjustments are done only to correct for omissions due to statistical reasons (for example Netherlands), while others include also the various elements of the hidden and illegal activities. The shares of unobserved activities for which the GDP has been corrected in the EU countries is documented in table 6.

A comparison of the share of adjustments for the inclusion of shadow activities in the

[9] There are also transitional supports available for regions that qualified during the previous period and for regions of the pre-2004 EU (EU 15) that would qualify if instead of the average of EU 25 (the EU after the eastward enlargement) the average of EU 15 would have been applied.

[10] §3.08 in the European System of National and Regional Accounts (ESA 1995) states that all productive activities, even if "they are illegal or not-registered at tax, social security, statistical and other public authorities", should be included in the production data.

Table 6: GDP adjustments for unobserved activities in the EU[a]

Country	Year[a]	Adjustment[a]	Year[b]	Adjustment[b]	Share of shadow economy[c]
Austria	2001	7.9			14.2
Belgium	2002	3.0–4.0[d]			16.9
Czech Republic	2000	4.6–9.3[d]	2001	8.4	17.1
Germany		n.a.			14.6
Denmark		n.i.			12.0
Estonia	2002	9.6	2001	7.4	16.6
Spain	2000	11.2			16.9
Finland	2002	n.a.			16.4
France		n.i.			15.8
Greece		n.i.			20.9
Hungary	2000	11.9	2001	16.0	22.4
Ireland	1998	4.0			15.9
Italy	2003	14.8–16.7[d]			20.5
Lithuania	2002	18.9	2001	18.3	22.4
Latvia	2000	8.28–13.6[d]	1998	16.8	20.4
Netherlands	1995	1.0			9.7
Poland	2002	7.8–15.7[d]	2001	14.3	29.2
Portugal		n.i.			19.3
Sweden	2000	1.3			15.9
Slovenia		n.i.	2001	6.7	16.6
Slovak Republic		n.i.	2001	14.5	22.3
United Kingdom	1996	0.9–2.6[e]			12.1

n.a. — not available. n.i. — not included in the survey.

[a] Source: United Nations (2008).

[b] Source: Feige & Urban (2008). Only for transition countries, based on correspondence with the respective statistical offices.

[c] Estimates from model M9 (spatially adjusted indicators, each with its own spatial effects coefficient).

[d] Different adjustments shares, depending on whether the production or expenditure based GDP calculation is considered.

[e] Only illegal activities.

GDP with our shadow economy estimates reveals that most of the countries obviously do not include shadow activities in a sufficient extent in their GDP calculation. In case of countries with a very low adjustment share (Ireland, the Netherlands, Sweden), there has been done almost no direct correction for shadow activities as reported in United Nations (2008). On the other extreme, for Italy and Lithuania the coverage of the direct adjustment is fairly similar to our shadow economy estimates (see table 6). However, as noted above, the true extent of shadow activities that is comprised in the GDP remains unknown.

Including shadow activities in the GDP in their full extent can change the eligibility of some EU regions for the financial support under the Convergence objective. In order to assess these effects, it is inevitable to utilise region specific shadow economy quotas such as derived in Section 4.2. For further processing, these estimates are corrected for the adjustment shares reported in table 6. Unfortunately, more recent data are unavailable, with the exception of Italy (see table 4). Also, being aware of possible double counting due to implicit coverage of shadow activities in the GDP, the shadow economy estimates are reduced by a third. The exact procedure of deriving the regional total GDP estimates is described in Appendix D.

As the poorest regions of the EU tend to have the highest shares of shadow economy (see figure 2), the inclusion or exclusion of shadow activities does not influence their eligibility for the financial support from the Eu budget. However, there are also several regions close to the critical 75% line of the EU average. Several of those would lose their eligibility if the shadow economic activity would be included in full extent in their GDP. Moreover, there are some regions that would become eligible. After the inclusion of the shadow economy also the EU average rises and in some countries shadow economy is already well captured in the GDP calculation.

The regions eligible for the financial resources under the Convergence objective are shown in figure 3.[11] The regions for which the eligibility status changes, depending on

[11] Our full listing of regions eligible under the Convergence objective differs slightly from the official listing in Commission Decision No 2006/595/EC (2006) due to the use of the GDP data for 2004. The

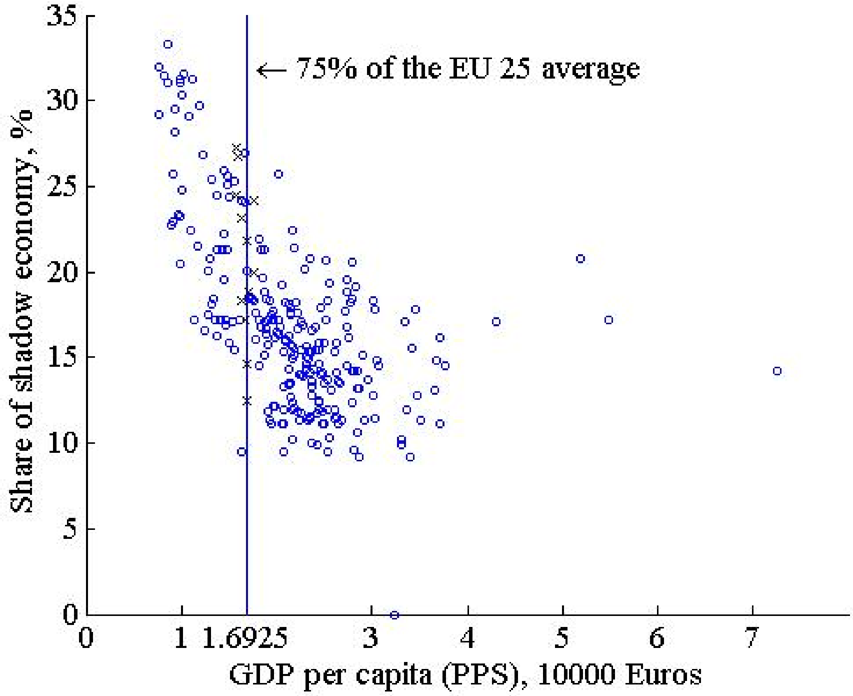


Figure 2: GDP per capita (PPS) and share of shadow economy in the EU NUTS 2 regions. Crosses mark the regions that traverse the 75% line if the official GDP is amended using our shadow economy estimates.

whether the official or estimated total GDP is used for determining the eligibility, are reported in table 7.

In general, the per capita GDP relative to the EU average increases the most in the regions of countries, which have a high share of shadow economy and that include—according to the available data—it only to a small extent in their GDP calculations. Thus, the increase in relative GDP per capita is largest in Greek regions. Conversely, the decrease in the share of GDP per capita in the EU average is largest for countries that capture the shadow activities in GDP to the full extent, which is found to be the case in Italy, Lithuania, Hungary, the Slovak Republic and Spain.

Some remarks are in order considering the listing of regions whose eligibility status changes after the inclusion of additional economic activity. First, for Greece there is no estimate available on the direct adjustments of GDP for the inclusion of shadow economy. If such adjustments are in fact done, then at least Thessalia should probably

official list is constructed using the 2001–2003 average figures.

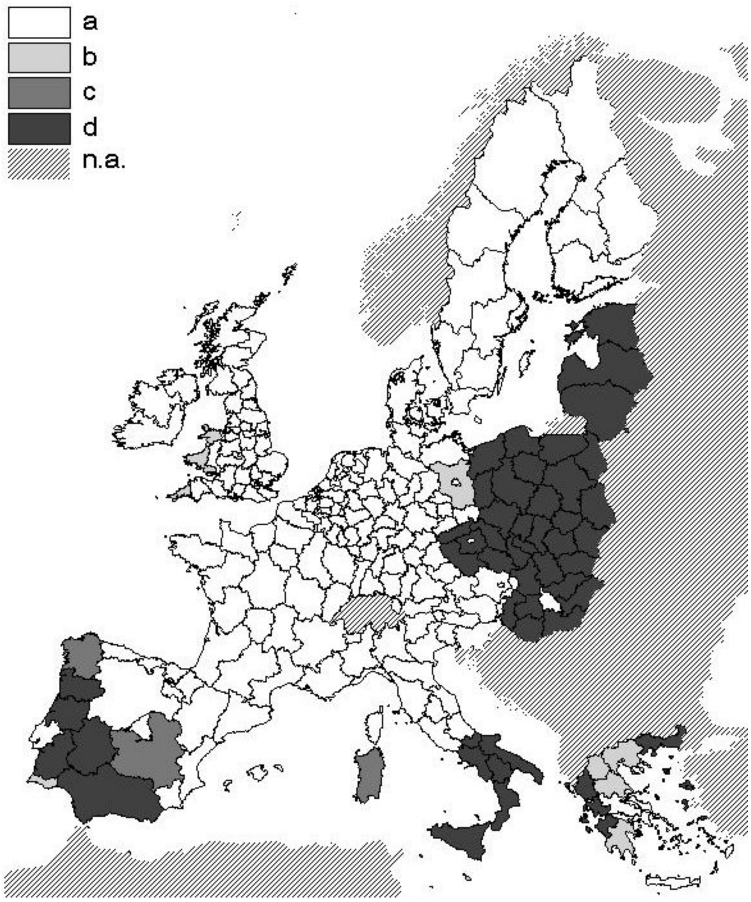


Figure 3: Eligibility of regions for financial resources under the Convergence objective. a—not eligible, b—eligible based on official GDP, c—eligible based on total GDP, d—eligible both based on official and total GDP, n.a.—not available.

Table 7: Eligibility under the Convergence objective

Region		Official GDP[a]	Total GDP[a]	Difference
PT15	Algarve	74.9	80.4	5.5
GR12	Kentriki Makedonia	72.7	78.5	5.8
GR25	Peloponnisos	70.9	78.0	7.1
MT00	Malta	74.0	77.5	3.5
GR13	Dytiki Makedonia	70.0	77.2	7.2
DE41	Brandenburg - Nordost	72.7	76.6	3.9
GR14	Thessalia	69.6	75.7	6.1
UKK3	Cornwall and Isles of Scilly	74.4	75.3	0.8
UKL1	West Wales and The Valleys	74.9	75.0	0.1
ES11	Galicia	78.0	73.7	-4.3
ITG2	Sardegna	78.0	73.7	-4.3
ES42	Castilla-la Mancha	75.8	71.6	-4.2

Source: Own calculations.

[a] As the share of the EU average.

still lie below the 75% line if total GDP is used to determine eligibility. A similar argument applies to regions in the UK listed in table 7 if the GDP includes shadow activities to a larger extent than assumed in the calculations. Second, Malta was out of the sample when estimating the shadow economy. It is unclear, how reliable it is to approximate its shadow economy quota with the EU average. Third, in Italy and Spain the statistical offices include a fairly large share of shadow economy in their GDP calculations, such that some of their regions do not qualify for the EU regional support based on the official GDP figures, but they would qualify if total GDP figures would be used to determine the eligibility. Thus, they cannot apply for the financial support from the EU budget under the Convergence objective due to good statistical work.

6 Conclusions

This paper provides for the first time estimates of shadow economy quotas in the regions of the EU at the NUTS 2 classification level (year 2004). To our knowledge, it is also the first contribution where MIMIC modeling is combined with some aspects of spatial econometrics. In addition, a new calibration procedure is proposed that is supposedly less sensitive with respect to the chosen base values than the existent approaches.

The estimation results reveal that the share of shadow economy exhibits considerable variation across the EU countries—confirming previous research—as well as across regions within countries. With a few exceptions, poor regions tend to have a higher share of shadow economy than rich ones.

In interpreting the results, it has to be kept in mind that there might be substantial measurement errors, though in the spatial models the estimated error of the latent variable is considerably smaller than in the non-spatial counterparts. Also, due to unavailability of data, some variables potentially affecting the size of shadow economy are omitted from the model. An example is here an unambiguous measure of tax paying moral. However, the country level estimates for the extent of shadow economy are fairly similar to the results from earlier studies. Also the ordering of the regions according to the extent of shadow economy and most of the shadow economy estimates are reasonable. The results are remarkably robust with respect to different model specifications and estimation techniques, including the consideration of spatial effects. Also, even though the estimates might show the potential of shadow activities rather than their true extent, policymakers should pay attention to their variation.

One of the main uses of the shadow economy estimates is assessing potentially lost tax revenues. An additional advantage of the regional estimates of the shadow economy is the possibility to capture this information in designing policies for combating shadow economy and stimulating development. In case of funding from the Structural Funds of the EU, our results suggest that it is important to include the effects to the shadow economy in deciding the financing of proposed projects. Moreover, it is necessary to achieve a good coverage of shadow economy in the official GDP figures in all countries in order to avoid unfair treatment of regions due to differences in the extent to which shadow economy is included in the GDP.

Specifically, there are regions whose eligibility for finances from the Funds depends on the extent to which shadow activities are captured in the GDP. Relying on total GDP that includes additional shadow activities as compared to the official GDP figures the eligibility status for such funding under the Convergence objective changes according to our calculations in the case of 12 regions. For example, according to the official figures the GDP per capita in Kentriki Makedonia and Peloponnisos reached respectively 72.7 and 70.9% of the EU average in 2004. As in Greece the GDP is adjusted only to a minor extent for the inclusion of shadow activities, these shares would rise to 78.5 and 78% if the whole shadow economy would be captured. In that case the regions would not be eligible for measures under the Convergence objective. For the Spanish regions Galicia and Castilla-la Mancha as well as for Sardegna in Italy the opposite is valid. In these countries shadow economy is well covered in the GDP calculations and the mentioned regions do not qualify for the convergence support from the EU. However, if all countries would include shadow activities to their full extent in the GDP, the mentioned regions would obtain eligibility.

Even if for some of the regions the total GDP may be overestimated, this result deserves attention of policymakers. In choosing the measures to be financed for achieving convergence it is important to consider whether they also help to transform shadow activities to the official economy, especially in case of the regions that are close to the official eligibility line or that benefit from transitional measures. This would improve the self-financing ability of the regions after the exclusion from the EU funding. Due to increased tax revenues the quality of the public services could be increased and, thus, the motivation to engage in the official sector instead of the shadow economy would rise further. Of course, such considerations are relevant also in case of the poorest regions.

The poorest regions are also those with the highest shares of shadow economy. These regions face the problem of meeting co-financing requirements both because of poverty and missed taxed revenues due to shadow economy. Thus, in order to help them catching up with the sustainable growth path, the co-financing requirements could be alleviated for those regions.

In general, policies for reducing shadow economy should be diversified across regions. It is necessary to ponder both the effects to the growth potential and to tax compliance. Especially in the poorest regions it might be rational to tolerate some excessive shadow economy for giving the people an additional source of income and keeping them in the labour market, even if hidden. This would also support the competitive position of these regions. Instead of punishing, rewarding mechanisms for the movement from the shadow to the official sector could be applied.

However, in designing the policies to combat shadow economy, it is important to consider both the short and long term effects of the implemented system of incentives. As discussed by Andreoni et al. (1998) the policy design usually relies only on economic incentives, though moral, psychological and institutional factors can also have an important impact on the decision of evading taxes. Moreover, according to Bowles (2008) and Bowles & Hwang (2008), economic incentives can have adverse effects on moral behaviour. Consistent with this, Enste (2008) suggests that in the long run it is ineluctable to build up strong tax moral, achieved by demonstrating that the tax revenues are used for delivering good public services. Therefore, though the model estimated in this paper suggests that the governments should pay attention to good labour market opportunities and keep taxes low in order to reduce participation in the shadow activities, it is also important to have an effective public sector and to encourage moral behaviour.

A Estimation of a MIMIC model

The model is solved for its reduced form by substituting the latent variable equation (1) into the measurement equations (2)

$$\mathbf{y}_r = \boldsymbol{\lambda}\boldsymbol{\gamma}'\mathbf{x}_r + \boldsymbol{\nu}_r, \quad \boldsymbol{\nu}_r = \boldsymbol{\lambda}\zeta_r + \boldsymbol{\varepsilon}_r. \tag{8}$$

The estimation of the model relies on fitting the implied covariance matrix of the observed variables to the sample covariance. The implied covariance matrix of the observed variables is $\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \mathbf{E}\left[\mathbf{z}_r\mathbf{z}_r'\right]$. The vector $\boldsymbol{\theta} = (\boldsymbol{\lambda}', \boldsymbol{\gamma}', \psi, \mathrm{vech}(\boldsymbol{\Phi})', \mathrm{vech}(\boldsymbol{\Theta}_\varepsilon)')'$ collects all model parameters[12] and $\mathbf{z}_r$ the observable variables (causes and indicators) of the model: $\mathbf{z}_r = (\mathbf{y}_r' \quad \mathbf{x}_r')'$. By nature of equation (8), the implied covariance matrix is

$$\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{\lambda}(\boldsymbol{\gamma}'\boldsymbol{\Phi}\boldsymbol{\gamma} + \psi)\boldsymbol{\lambda}' + \boldsymbol{\Theta}_\varepsilon & \boldsymbol{\lambda}\boldsymbol{\gamma}'\boldsymbol{\Phi} \\ \boldsymbol{\Phi}\boldsymbol{\gamma}\boldsymbol{\lambda}' & \boldsymbol{\Phi} \end{bmatrix}. \tag{9}$$

The sample covariance matrix $\mathbf{S}$ is

$$\mathbf{S} = \begin{bmatrix} \mathbf{E}[\mathbf{y}_r\mathbf{y}_r'] & \mathbf{E}[\mathbf{y}_r\mathbf{x}_r'] \\ \mathbf{E}[\mathbf{x}_r\mathbf{y}_r'] & \mathbf{E}[\mathbf{x}_r\mathbf{x}_r'] \end{bmatrix}. \tag{10}$$

[12] The operator vech denotes half-vectorisation, i.e. it transforms the matrix into a vector that collects the unique elements of the matrix.

The covariance matrix of the cause variables is fixed to be equal to the sample counterpart, i.e. $\mathbf{\Phi} = \mathbf{E}[\mathbf{x}_r\mathbf{x}_r']$. Still, the rest of the parameters is not identified. To identify the parameters, one of the elements in $\boldsymbol{\lambda}$ has to be fixed. Then, there remain $p+q+1/2p(p+1)$ free parameters that have to be estimated based on $1/2(p+q)(p+q+1)$ equations.

The model parameters $\boldsymbol{\theta}$ are estimated by minimizing a distance measure with respect to $\boldsymbol{\theta}$. Common distance criteria, Unweighted Least Squares (ULS), Generalized Least Squares (GLS) and Maximum Likelihood (ML), are formalized as follows:

$$F_{ULS} = \left(\frac{1}{2}\right) \operatorname{tr}\left\{[\mathbf{S} - \mathbf{\Sigma}(\boldsymbol{\theta})]^2\right\}, \tag{11}$$

$$F_{GLS} = \left(\frac{1}{2}\right) \operatorname{tr}\left\{\left[\mathbf{I} - \mathbf{\Sigma}(\boldsymbol{\theta})\mathbf{S}^{-1}\right]^2\right\}, \tag{12}$$

$$F_{ML} = \ln|\mathbf{\Sigma}(\boldsymbol{\theta})| + \operatorname{tr}\left\{\mathbf{S}\mathbf{\Sigma}^{-1}(\boldsymbol{\theta})\right\} - \log|\mathbf{S}| - (p+q). \tag{13}$$

The ML objective function is derived under the presumption of a multivariate normal distribution. In case of the GLS criterion, the underlying assumptions are less restrictive. The observations have to be i.i.d. with any distribution characterised by moderate kurtosis (Bollen 1989). Both methods are suitable for inferential purposes if the distributional assumptions are satisfied. The ULS objective function is not derived from a distributional assumption. Thus, inferential statistics cannot be obtained directly (Backhaus et al. 2006), but could be derived by means of resampling methods instead.

B Adjusting for spatial effects

In models 8, 9 and 10 it is assumed that the indicators or causes are measured with an error due to spatial interdependencies. Thus, instead of the original variables $\mathbf{z}_i$ (a $R \times 1$ vector of the R observations of the ith observed variable, $i \in \{1, 2, \ldots, p+q\}$), the variables $\tilde{\mathbf{z}}_i = (\mathbf{I} - \rho_i\mathbf{W})\mathbf{z}_i$ are used in the models, with $\tilde{\mathbf{z}}_i$ denoting the vector of spatially adjusted observations for the ith observed variable, ρ_i the spatial effects coefficient corresponding to the ith variable and $\mathbf{W}$ the spatial weighting matrix. In the following sections the derivation of the weighting matrix and the procedure of choosing an appropriate parametrization of the spatial effects are described.

B.1 The weighting matrix

The weighting matrix is based on a distance matrix, derived from a distance matrix for NUTS 3 regions.[13] In order to obtain the distances between the NUTS 2 regions, the land area shares of NUTS 3 regions within each NUTS 2 region are calculated. In case of missing data, those are replaced with A^{ud}/N_r, where N_r is the number of NUTS 3 regions in a NUTS 2 region r and A_r^{ud} denotes the share of the NUTS 2 area for which the area data at the NUTS 3 level are missing ("undefined area"). Thus, $A_r^{ud} = 1$ if for none of the NUTS 3 regions within a NUTS 2 region r area data are available. If for some NUTS 3 regions data are available but for some others within the same NUTS 2 region not, $A_r^{ud} = 1 - A_r^d$, with A_r^d denoting the share of the NUTS 2 region r's area for which area data at NUTS 3 level are available.

The distances d_{rs} between the NUTS 2 regions r and s are calculated as

$$d_{rs} = \begin{cases} \sum_{k \in r} \sum_{l \in s} a_{rk} a_{sl} d_{kl} & \text{if } r \neq s \\ 0 & \text{if } r = s \end{cases}, \tag{14}$$

where a_{rk} and a_{sl} denote the land area share of a NUTS 3 region k in a NUTS 2 region r and the land area share of a NUTS 3 region l in a NUTS 2 region s, respectively. The original distance between the NUTS 3 regions k and l is denoted by d_{kl}.

In order to convert the distance matrix into a weighting matrix, the so-called tri-cube function has been used. This function has also been used e.g. by McMillen (1996) and

[13] We are grateful to Artem Korzhenevych for providing this matrix.

is formally

$$w_{rs} = \begin{cases} \left(1 - \left(\frac{d_{rs}}{D}\right)^3\right)^3 & \text{if } d_{rs} < D \text{ and } r \neq s \\ 0 & \text{if } d_{rs} \geq D \text{ or } r = s. \end{cases} \tag{15}$$

Notably the tri-cube approach implies that the region s is considered to have a spatial effect on region r ($w_{rs} > 0$) only if its distance from region r is not longer than D kilometers. Therefore, the number of regions exerting a spatial effect varies across the regions.[14] In the estimation the weighting matrix $\mathbf{W}$ has been row-standardized, such that $\sum_s w_{rs} = 1$.

B.2 Spatial effects in the SEM

Three sets of models have been estimated.

(i) Assuming that all variables (both indicators and causes) should be adjusted for spatial effects.

(ii) Assuming that only the indicator variables ($\mathbf{y}_r$) should be adjusted for spatial effects.

(iii) Assuming that only the cause variables ($\mathbf{x}_r$) should be adjusted for spatial effects.

In all of the cases, it is initially assumed that for all of the spatially affected variables the same spatial effects coefficient can be used for the adjustment, i.e. an identical ρ is used for all of the variables in the transformation: $\tilde{\mathbf{z}}_i = (\mathbf{I} - \rho\mathbf{W})\mathbf{z}_i$. After the spatial adjustment the data are transformed as the non-weighted data in the unweighted model: all variables are centered and divided by the mean. Then the MIMIC estimation is applied.

In choosing the most appropriate model, the coefficient of determination for the latent variable (calculated as $R^2_\eta = 1 - \psi/\operatorname{Var}[\eta]$) is used as the selection criterion.[15]

[14] An alternative approach would be to fix the number of closest regions that are considered to have a spatial effect to each region. As the geographical size of the NUTS 2 regions varies a lot, we believe that the distance approach is more valid for characterising the spatial extent of the spatial effects.

[15] Using the general goodness of fit index or the value of the objective function are not appropriate as the sample covariance matrix depends on the data transformation and, thus, these measures are not

The procedure of choosing the cut-off distance D and the spatial effects coefficient ρ is found in following steps.

1. R^2_η is maximized both with respect to ρ and D.[16] The cut-off distance D is assumed to be in the range 30 km $\leq D \leq$ 150 km (step size 10 km), and the spatial effects coefficient in the range $-1 \leq \rho \leq 1$, with step size 0.1.

2. The grid for the cut-off distance is reduced to 5 km around the obtained maximum from step 1. Again, the model with the maximum R^2_η is chosen. The corresponding D is taken as the optimal cut-off distance (D^*).

3. Step size 0.02 is chosen for ρ in the neighbourhood of the ρ obtained in step 2 and R^2_η is maximised over this grid to find the most appropriate ρ^*.

The most appropriate model is considered to be the one that uses the weighting matrix with cut-off distance obtained in step 2 and the spatial effects coefficient from step 3.

In case of model 9 in table 2, the two indicators are assumed to have distinct spatial effects coefficients. It has been assumed that the cut-off distance is in both cases as in the model with identical spatial effects coefficient (model 8). Then, the model has been estimated for all pairwise combinations of ρ from $-0.5 < \rho < 0.5$, with step size 0.05. Again, the decision criterion is R^2_η. After that we refine the grid to 0.02 in the surrounding of the initial values of the ρs. In all of the models with GLS estimation, the ULS estimates have been taken as the starting values. For ULS, the starting values are those from the unweighted model.

In table 2 the model with all variables being weighted according to the same weighting scheme is not presented due to unreasonable parameter estimates in the model chosen by the procedure described above. In case of models 8 and 9 (indicators spatially adjusted, model 8 using the same ρ and model 9 distinct values of ρ for each indicator), the cut-off distance D^* is 90 km. This appears reasonable, considering that it could be interpreted

comparable across models estimated with differently transformed data.

[16]In order to avoid the problem of random peaks, in fact the moving average of R^2_η across ρs with step size 3 is maximised. The same procedure is applied in step 2.

as a commuting distance. In model 8 the spatial effects coefficient $\rho^* = -0.06$. In model 9 $\rho^* = 0.00$ in case of GDP per capita and $\rho^* = -0.18$ for labour force participation rate. In model 10 where only the cause variables are spatially adjusted, the cut-off distance D^* is 50 km and the spatial effects coefficient $\rho^* = 0.10$.

C Shadow economy in the EU NUTS 2 regions

Table 8: Shadow economy quotas (% in GDP) in the EU NUTS 2 regions in 2004[a]

Country	Region	Region's name	M1	M1[b]	M1[c]	M9[d]	M10[e]
AT	AT31	Oberösterreich	13.0	13.9	14.5	13.6	13.3
AT	AT34	Vorarlberg	13.6	14.8	14.9	14.2	13.9
AT	AT12	Niederösterreich	13.7	13.7	14.9	14.0	13.9
AT	AT22	Steiermark	13.9	13.7	15.1	14.1	14.1
AT	AT33	Tirol	13.9	14.1	15.1	14.2	14.2
AT	AT13	Wien	14.2	15.1	15.3	14.5	14.4
AT	AT21	Kärnten	14.3	14.4	15.4	14.6	14.5
AT	AT32	Salzburg	14.4	13.9	15.4	14.5	14.5
AT	AT11	Burgenland (A)	15.1	15.1	15.9	15.2	15.3
BE	BE21	Prov. Antwerpen	14.8	13.6	15.7	14.8	13.2
BE	BE24	Prov. Vlaams Brabant	15.0	12.3	15.8	14.5	12.7
BE	BE23	Prov. Oost-Vlaanderen	15.2	13.0	15.9	14.9	12.9
BE	BE25	Prov. West-Vlaanderen	16.3	13.4	16.6	15.8	16.1
BE	BE22	Prov. Limburg (B)	16.4	13.9	16.7	15.8	15.2
BE	BE31	Prov. Brabant Wallon	17.9	14.1	17.7	17.2	15.6
BE	BE34	Prov. Luxembourg (B)	18.3	15.5	17.9	17.6	18.1
BE	BE35	Prov. Namur	18.5	14.4	18.1	17.2	18.2
BE	BE33	Prov. Liège	19.1	16.7	18.4	18.3	18.9
BE	BE32	Prov. Hainaut	19.3	17.0	18.5	18.4	19.0
BE	BE10	Région de Bruxelles-Capitale/Brussels Hoofdstedelijk Gewest	21.5	19.5	20.0	20.8	19.5
CZ	CZ03	Jihozápad	15.4	16.1	16.1	15.9	15.7
CZ	CZ05	Severovýchod	15.7	16.9	16.2	16.3	15.9
CZ	CZ06	Jihovýchod	16.2	17.5	16.6	16.8	16.5
CZ	CZ02	Strední Cechy	16.7	16.7	16.9	17.1	15.0
CZ	CZ07	Strední Morava	16.7	18.7	16.9	17.5	17.0
CZ	CZ04	Severozápad	17.2	20.0	17.2	18.1	17.6
CZ	CZ08	Moravskoslezsko	17.5	20.8	17.4	18.4	17.8
CZ	CZ01	Praha	17.5	15.0	17.4	17.1	15.6
DE	DE11	Stuttgart	12.8	13.3	14.4	12.8	13.0
DE	DE71	Darmstadt	13.0	12.6	14.5	12.8	13.1
DE	DEB3	Rheinhessen-Pfalz	13.6	12.0	14.9	12.8	13.6
DE	DE12	Karlsruhe	13.6	13.1	14.9	13.2	13.8
DE	DE21	Oberbayern	13.7	11.8	15.0	13.1	13.7
DE	DE23	Oberpfalz	13.8	12.8	15.0	13.1	13.9
DE	DE91	Braunschweig	13.9	13.7	15.1	13.3	14.0
DE	DE25	Mittelfranken	14.0	13.9	15.2	13.7	14.2
DE	DE22	Niederbayern	14.1	13.3	15.2	13.5	14.2
DE	DE14	Tübingen	14.1	13.2	15.2	13.5	14.2

Country	Region	Region's name	M1	M1[b]	M1[c]	M9[d]	M10[e]
DE	DEC0	Saarland	14.3	13.5	15.3	13.6	14.3
DE	DE26	Unterfranken	14.3	13.6	15.3	13.7	14.4
DE	DEB1	Koblenz	14.4	12.8	15.4	13.5	14.5
DE	DEB2	Trier	14.5	12.3	15.5	13.4	14.5
DE	DE27	Schwaben	14.6	13.8	15.5	14.1	14.7
DE	DE13	Freiburg	14.6	13.5	15.6	14.0	14.7
DE	DEF0	Schleswig-Holstein	14.8	13.2	15.7	14.0	14.8
DE	DEA1	Düsseldorf	14.8	14.4	15.7	14.2	14.9
DE	DE24	Oberfranken	14.8	15.2	15.7	14.6	15.0
DE	DE92	Hannover	15.0	14.0	15.8	14.3	15.1
DE	DE73	Kassel	15.2	13.2	15.9	14.1	15.2
DE	DEA2	Köln	15.2	13.0	15.9	14.1	15.2
DE	DE93	Lüneburg	15.3	14.0	16.0	14.6	15.3
DE	DE72	Gieen	15.3	13.6	16.0	14.3	15.3
DE	DEA3	Münster	15.4	13.8	16.1	14.4	15.4
DE	DEA5	Arnsberg	15.6	15.5	16.2	15.0	15.8
DE	DE94	Weser-Ems	16.1	14.4	16.5	15.1	16.1
DE	DEA4	Detmold	16.1	15.4	16.5	15.4	16.2
DE	DEG0	Thüringen	16.4	17.3	16.7	16.1	16.5
DE	DE42	Brandenburg - Südwest	17.5	17.9	17.4	17.1	17.6
DE	DE60	Hamburg	17.9	16.6	17.7	17.1	17.9
DE	DED2	Dresden	18.0	19.2	17.7	17.7	18.1
DE	DED1	Chemnitz	18.4	20.4	18.0	18.4	18.6
DE	DEE0	Sachsen-Anhalt	18.5	20.5	18.1	18.3	18.7
DE	DED3	Leipzig	18.6	20.0	18.1	18.4	18.7
DE	DE41	Brandenburg - Nordost	18.7	19.9	18.2	18.3	18.8
DE	DE30	Berlin	18.7	17.7	18.2	17.9	18.6
DE	DE80	Mecklenburg-Vorpommern	18.8	20.4	18.2	18.5	18.9
DE	DE50	Bremen	18.8	18.0	18.3	17.8	18.8
DK	DK00	Denmark	10.8	12.5	13.1	12.0	11.1
EE	EE00	Estonia	16.3	18.7	16.6	16.6	16.7
ES	ES30	Comunidad de Madrid	13.1	15.0	14.6	13.2	13.6
ES	ES24	Aragón	15.1	16.0	15.9	15.3	15.5
ES	ES51	Cataluña	15.2	18.0	15.9	15.8	15.7
ES	ES62	Región de Murcia	15.9	19.5	16.4	16.7	16.5
ES	ES12	Principado de Asturias	16.0	17.9	16.4	16.4	16.4
ES	ES23	La Rioja	16.1	17.3	16.5	16.6	16.4
ES	ES22	Comunidad Foral de Navarra	16.1	16.7	16.5	16.2	16.4
ES	ES52	Comunidad Valenciana	16.1	19.2	16.5	16.8	16.6
ES	ES21	Pais Vasco	16.5	18.5	16.8	16.8	16.9
ES	ES13	Cantabria	17.6	19.5	17.5	18.1	18.0
ES	ES53	Illes Balears	17.6	19.1	17.5	17.9	18.0

Country	Region	Region's name	M1	M1[b]	M1[c]	M9[d]	M10[e]
ES	ES42	Castilla-la Mancha	18.7	19.5	18.2	18.8	18.9
ES	ES41	Castilla y León	19.2	19.9	18.5	19.3	19.4
ES	ES11	Galicia	19.3	21.4	18.6	19.9	19.6
ES	ES61	Andalucia	19.6	22.9	18.8	20.1	20.0
ES	ES43	Extremadura	22.1	23.8	20.4	22.2	22.4
FI	FI20	Åland	1.8	0.6	7.3	-0.1	2.1
FI	FI18	Etelä-Suomi	14.5	15.9	15.5	15.2	14.8
FI	FI19	Länsi-Suomi	17.2	17.9	17.2	17.6	17.3
FI	FI1A	Pohjois-Suomi	17.9	18.4	17.7	18.2	18.0
FI	FI13	Itä-Suomi	19.6	19.7	18.7	19.6	19.6
FR	FR42	Alsace	13.8	14.9	15.0	13.9	14.1
FR	FR10	Île de France	15.1	15.3	15.8	14.8	15.2
FR	FR23	Haute-Normandie	15.2	15.6	15.9	14.9	15.4
FR	FR43	Franche-Comté	15.6	15.8	16.2	15.4	15.8
FR	FR71	Rhône-Alpes	15.6	15.8	16.2	15.4	15.8
FR	FR24	Centre	15.8	15.0	16.3	15.3	15.9
FR	FR22	Picardie	15.9	16.8	16.3	15.8	16.1
FR	FR51	Pays de la Loire	15.9	15.5	16.4	15.6	16.0
FR	FR26	Bourgogne	16.4	15.9	16.7	16.0	16.5
FR	FR41	Lorraine	16.4	17.2	16.7	16.1	16.6
FR	FR52	Bretagne	16.5	15.3	16.8	16.1	16.6
FR	FR72	Auvergne	16.7	16.0	16.9	16.5	16.8
FR	FR25	Basse-Normandie	16.7	15.5	16.9	16.2	16.8
FR	FR53	Poitou-Charentes	16.8	16.2	17.0	16.4	16.9
FR	FR62	Midi-Pyrénées	16.8	15.3	17.0	16.3	16.9
FR	FR30	Nord - Pas-de-Calais	16.9	17.8	17.0	16.7	17.1
FR	FR21	Champagne-Ardenne	17.2	16.6	17.2	16.7	17.3
FR	FR82	Provence-Alpes-Cô te d'Azur	17.8	16.4	17.6	17.1	17.8
FR	FR63	Limousin	18.2	16.1	17.8	17.5	18.1
FR	FR61	Aquitaine	18.3	16.7	17.9	17.6	18.3
FR	FR81	Languedoc-Roussillon	20.0	17.2	19.0	18.8	19.9
FR	FR83	Corse	24.9	17.0	22.1	21.3	24.3
GR	GR22	Ionia Nisia	5.9	10.2	10.0	9.5	6.3
GR	GR42	Notio Aigaio	18.0	16.6	17.7	18.3	18.0
GR	GR30	Attiki	18.1	18.4	17.8	18.2	18.3
GR	GR41	Voreio Aigaio	20.1	16.0	19.1	19.5	19.9
GR	GR43	Kriti	21.5	19.5	20.0	21.9	21.5
GR	GR12	Kentriki Makedonia	22.8	22.4	20.8	23.2	22.8
GR	GR11	Anatoliki Makedonia, Thraki	23.9	23.2	21.5	24.5	23.9
GR	GR14	Thessalia	24.4	22.5	21.8	24.5	24.4
GR	GR24	Sterea Ellada	25.1	25.0	22.3	25.7	25.1
GR	GR23	Dytiki Ellada	25.1	24.0	22.3	25.4	25.1

Country	Region	Region's name	M1	M1^b	M1^c	M9^d	M10^e
GR	GR21	Ipeiros	25.4	22.9	22.5	25.3	25.3
GR	GR13	Dytiki Makedonia	26.6	26.5	23.3	27.2	26.6
GR	GR25	Peloponnisos	26.7	23.7	23.3	26.8	26.6
HU	HU21	Közép-Dunántúl	20.1	20.8	19.1	20.8	20.3
HU	HU22	Nyugat-Dunántúl	20.9	20.9	19.6	21.3	21.0
HU	HU10	Közép-Magyarország	22.7	20.5	20.7	22.4	22.6
HU	HU31	Észak-Magyarország	22.7	23.1	20.7	23.0	22.7
HU	HU32	Észak-Alföld	22.7	21.8	20.8	22.7	22.7
HU	HU33	Dél-Alföld	23.4	21.9	21.2	23.4	23.3
HU	HU23	Dél-Dunántúl	23.4	22.2	21.2	23.3	23.3
IE	IE02	Southern and Eastern	13.9	18.8	15.1	15.6	14.7
IE	IE01	Border, Midlands and Western	16.0	19.7	16.5	17.5	16.7
IT	ITC4	Lombardia	17.1	16.6	17.2	17.8	17.2
IT	ITD2	Provincia Autonoma Trento	18.1	15.1	17.8	17.7	18.0
IT	ITD4	Friuli-Venezia Giulia	18.1	16.3	17.8	18.3	18.1
IT	ITD3	Veneto	18.3	17.4	17.9	18.9	18.4
IT	ITD1	Provincia Autonoma Bolzano-Bozen	18.6	16.0	18.1	18.3	18.5
IT	ITD5	Emilia-Romagna	18.8	17.0	18.2	19.2	18.8
IT	ITC1	Piemonte	18.9	17.7	18.3	19.4	19.0
IT	ITE3	Marche	20.0	18.2	19.0	20.2	20.0
IT	ITC2	Valle d'Aosta/Vallée d'Aoste	20.5	15.6	19.3	19.6	20.3
IT	ITE1	Toscana	20.6	18.1	19.4	20.7	20.5
IT	ITC3	Liguria	21.2	17.6	19.8	20.7	21.0
IT	ITE4	Lazio	21.2	17.8	19.8	20.6	21.0
IT	ITF1	Abruzzo	21.5	18.8	20.0	21.3	21.3
IT	ITE2	Umbria	21.9	18.2	20.2	21.4	21.7
IT	ITF4	Puglia	24.2	23.8	21.7	24.4	24.1
IT	ITG2	Sardegna	24.2	22.9	21.7	24.1	24.1
IT	ITF5	Basilicata	24.3	22.7	21.8	24.2	24.2
IT	ITF2	Molise	24.5	21.6	21.9	24.0	24.3
IT	ITF6	Calabria	25.8	23.4	22.7	25.1	25.6
IT	ITF3	Campania	26.0	24.3	22.9	25.6	25.8
IT	ITG1	Sicilia	26.7	24.6	23.3	26.0	26.5
LT	LT00	Lithuania	22.9	22.8	20.9	22.4	23.1
LV	LV00	Latvia	21.0	21.6	19.6	20.4	21.2
NL	NL23	Flevoland	8.6	10.6	11.7	9.5	9.0
NL	NL33	Zuid-Holland	8.8	9.4	11.8	9.2	9.0
NL	NL42	Limburg (NL)	8.8	10.0	11.8	9.5	7.3
NL	NL41	Noord-Brabant	8.9	10.1	11.9	9.6	9.2
NL	NL31	Utrecht	9.0	8.7	11.9	9.2	9.2

Country	Region	Region's name	M1	M1[b]	M1[c]	M9[d]	M10[e]
NL	NL21	Overijssel	9.4	10.2	12.2	10.0	9.7
NL	NL32	Noord-Holland	9.5	9.4	12.2	9.9	9.7
NL	NL22	Gelderland	9.7	9.9	12.4	10.0	9.9
NL	NL13	Drenthe	9.8	10.4	12.5	10.3	8.9
NL	NL34	Zeeland	9.9	10.1	12.5	10.3	8.5
NL	NL11	Groningen	10.0	9.8	12.6	10.2	9.0
NL	NL12	Friesland (NL)	11.1	10.8	13.3	11.3	11.3
PL	PL22	Slaskie	26.2	29.7	23.0	26.9	26.4
PL	PL12	Mazowieckie	26.7	26.7	23.3	26.9	26.7
PL	PL52	Opolskie	27.7	29.7	23.9	28.2	27.8
PL	PL63	Pomorskie	28.8	30.8	24.6	29.1	28.8
PL	PL32	Podkarpackie	28.8	29.1	24.6	29.2	28.7
PL	PL21	Malopolskie	29.0	29.8	24.8	29.5	29.0
PL	PL41	Wielkopolskie	29.2	31.0	24.9	29.7	29.2
PL	PL11	Lódzkie	29.7	31.2	25.3	30.3	29.8
PL	PL51	Dolnoslaskie	30.7	33.8	25.9	31.3	30.8
PL	PL62	Warminsko-Mazurskie	30.8	32.8	26.0	31.1	30.9
PL	PL43	Lubuskie	30.9	33.1	26.0	31.1	30.9
PL	PL61	Kujawsko-Pomorskie	30.9	33.0	26.0	31.3	30.9
PL	PL42	Zachodniopomorskie	31.1	34.1	26.1	31.5	31.2
PL	PL34	Podlaskie	31.2	30.1	26.2	31.5	31.0
PL	PL31	Lubelskie	31.7	30.2	26.5	31.9	31.4
PL	PL33	Swietokrzyskie	32.9	32.8	27.3	33.3	32.8
PT	PT17	Lisboa	16.7	17.8	16.9	16.9	17.0
PT	PT11	Norte	19.0	20.9	18.3	20.1	19.3
PT	PT16	Centro (PT)	20.8	19.5	19.5	21.3	20.9
PT	PT18	Alentejo	21.5	21.4	19.9	21.3	21.7
PT	PT15	Algarve	22.0	20.5	20.3	21.8	22.1
SE	SE23	Västsverige	15.2	14.2	15.9	15.4	15.1
SE	SE31	Norra Mellansverige	15.3	14.0	16.0	15.4	15.2
SE	SE21	Småland med öarna	15.3	13.5	16.0	15.4	15.2
SE	SE12	Östra Mellansverige	15.6	14.4	16.1	15.6	15.5
SE	SE33	Övre Norrland	15.8	13.3	16.3	15.4	15.6
SE	SE32	Mellersta Norrland	16.1	13.6	16.5	15.8	15.9
SE	SE11	Stockholm	16.1	14.2	16.5	16.2	16.0
SE	SE22	Sydsverige	16.7	15.3	16.9	16.8	16.6
SI	SI00	Slovenia	16.2	17.7	16.6	16.6	16.5
SK	SK01	Bratislavský kraj	18.7	18.7	18.2	18.4	18.9
SK	SK02	Západné Slovensko	20.9	24.2	19.6	21.5	21.3
SK	SK03	Stredné Slovensko	24.3	28.3	21.8	24.8	24.6
SK	SK04	Východné Slovensko	25.1	29.8	22.3	25.7	25.4
UK	UKD2	Cheshire	10.1	12.1	12.6	10.6	9.3
UK	UKJ1	Berkshire, Bucks and Oxfordshire	10.5	12.7	12.9	11.1	11.1

Country	Region	Region's name	M1	M1[b]	M1[c]	M9[d]	M10[e]
UK	UKF2	Leicestershire, Rutland and Northants	10.6	12.9	13.0	11.1	11.1
UK	UKE4	West Yorkshire	10.6	13.2	13.0	11.1	9.9
UK	UKG2	Shropshire and Staffordshire	10.7	12.9	13.1	11.2	11.3
UK	UKH2	Bedfordshire, Hertfordshire	10.9	12.5	13.1	11.3	11.4
UK	UKC1	Tees Valley and Durham	10.9	13.3	13.1	11.1	11.4
UK	UKD3	Greater Manchester	10.9	12.8	13.2	11.1	10.1
UK	UKH3	Essex	10.9	13.3	13.2	11.7	11.5
UK	UKE3	South Yorkshire	11.0	12.8	13.2	11.2	10.3
UK	UKF1	Derbyshire and Nottinghamshire	11.0	12.8	13.2	11.3	10.2
UK	UKK1	Gloucestershire, Wiltshire and Bristol/Bath area	11.2	12.6	13.3	11.4	11.6
UK	UKG1	Herefordshire, Worcestershire and Warks	11.2	12.6	13.3	11.5	10.3
UK	UKG3	West Midlands	11.2	14.4	13.3	11.8	10.5
UK	UKM2	Eastern Scotland	11.2	12.7	13.4	11.3	11.7
UK	UKJ3	Hampshire and Isle of Wight	11.3	12.5	13.4	11.4	11.7
UK	UKM3	South Western Scotland	11.4	13.6	13.4	11.6	11.9
UK	UKL2	East Wales	11.5	12.5	13.5	11.6	11.9
UK	UKD1	Cumbria	11.5	13.6	13.5	12.1	12.0
UK	UKH1	East Anglia	11.7	12.7	13.7	11.8	12.2
UK	UKE1	East Yorkshire and Northern Lincolnshire	11.8	13.8	13.7	12.0	12.3
UK	UKF3	Lincolnshire	11.8	13.8	13.7	12.2	12.3
UK	UKD4	Lancashire	12.0	12.8	13.9	11.9	11.1
UK	UKD5	Merseyside	12.0	13.1	13.9	11.8	11.2
UK	UKC2	Northumberland, Tyne and Wear	12.1	13.3	13.9	11.9	12.5
UK	UKJ4	Kent	12.1	13.3	13.9	12.4	12.5
UK	UKI2	Outer London	12.1	13.7	13.9	12.5	11.0
UK	UKJ2	Surrey, East and West Sussex	12.3	12.7	14.0	12.4	12.6
UK	UKL1	West Wales and The Valleys	12.5	13.5	14.2	12.4	12.9
UK	UKE2	North Yorkshire	12.7	12.1	14.3	12.4	13.0
UK	UKK2	Dorset and Somerset	12.9	12.6	14.4	12.7	13.2
UK	UKK4	Devon	13.5	13.3	14.8	13.3	13.8
UK	UKN0	Northern Ireland	13.8	13.6	15.0	13.5	14.1

Country	Region	Region's name	M1	M1[b]	M1[c]	M9[d]	M10[e]
UK	UKI1	Inner London	14.0	15.4	15.1	14.2	12.9
UK	UKK3	Cornwall and Isles of Scilly	14.8	14.4	15.7	14.7	15.1

Source: Own calculations.

[a]Based on GLS estimates, if not indicated otherwise. Based on the weighted average and variance, the regional GDP shares have been used as the weights. The calibration procedure using average = 17.2 and standard deviation = 5.4 has been applied if not indicated otherwise.

[b]Based on ULS estimates.

[c]The calibration procedure using only average = 17.2.

[d]Indicators spatially adjusted, each with its own spatial effects coefficient.

[e]Cause variables spatially adjusted.

D Calculating regional total GDPs

In deriving total GDP from the official GDP figures, several aspects have to be considered. First, the official GDP includes direct adjustments for shadow activities as reported in table 6. Second, parts of shadow economy have also been included implicitly. That is, the official GDP in region r (Y_r^o) can be considered to consist of three components:

$$Y_r^o = Y_r^f + Y_r^{si} + Y_r^{sd}, \tag{16}$$

where Y_r^f stands for GDP from the formal economy (the part of economy that does not hide its activities from various governmental authorities), Y_r^{si} for implicitly in GDP included shadow economy and Y_r^{sd} for GDP corresponding to directly included shadow economy. The third aspect to be considered is the base of the shadow economy estimates. We regard our estimates to be measured with respect to the formal economy. Thus, the total GDP including both formal and shadow economy can be calculated as

$$Y_r^t = (1 + SE_r)Y_r^f, \tag{17}$$

where SE_r denotes our estimate of the shadow economy in region r.

However, Y_r^f is unknown. In order to calculate it, it is necessary to subtract from the official GDP figures both the directly and implicitly included shadow economy. As the latter is unavailable, it is assumed that a third of our shadow economy estimates is included implicitly in the official GDP data. Formally, $Y_r^{si}/Y_r^f = 1/3SE_r$. If the share of directly included shadow activities $d_r = Y_r^{sd}/Y_r^o$ were available for each region, the total GDP could be calculated as

$$Y_r^t = (1 + SE_r)\frac{(1 - d_r)}{(1 + \frac{1}{3}SE_r)}Y_r^o. \tag{18}$$

Though, the data on direct adjustments of GDP for comprising shadow activities are available only at country level as reported in table 6.[17] In order to overcome this problem, we assume that the share of shadow economy b_r that is directly included in GDP is constant across the regions of each country i and corresponds to that of the country:

$$b_{r\in i} = \frac{Y_i^{sd}/Y_i^f}{SE_i} = \frac{d_i}{1 - d_i}\frac{(1 + \frac{1}{3}SE_i)}{SE_i}, \quad b_{r\in i} = b_i. \tag{19}$$

Substituting in equation (19) country variables for regional variables (except for b_i) and solving the result for $d_{r\in i}$ gives

$$d_{r\in i} = \frac{b_i SE_r}{1 + b_i SE_r + \frac{1}{3}SE_r}. \tag{20}$$

Combining the result with equation (18) enables to calculate regional total GDPs.

In case of countries for which no data are available on the shadow economy adjustments of GDP, the direct adjustments are assumed to be zero. For Italy and Lithuania we assume that the implicit and direct adjustments fully cover the shadow activities, such that official GDP is equal to total GDP. Further calculations revealed that also in Hungary, Slovak Republic and Spain shadow activities are covered by the official GDP to a sufficiently large extent. For the regions of UK that were omitted from the sample due to unavailability of data, the share of shadow economy is assumed to equal that of

[17] In case of countries with several adjustment shares, the average of them is used. For the Eastern and Central European countries either the data from United Nations (2008) or Feige & Urban (2008) are used, depending on which source delivers later data.

the closest region. For the rest of the countries and regions omitted from the sample the EU average shadow economy (17.2%) has been applied.

References

Andreoni, J., Erard, B. & Feinstein, J. (1998), 'Tax compliance', *Journal of Economic Literature* **36**(2), 818–860.

Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2006), *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung*, 11 edn, Springer, Berlin, Heidelberg, New York.

Bernotaitė, R. & Piskunova, A. (2005), An expenditure-based estimate of Latvia's shadow economy. SSE Riga Working Papers 2005:5 (73). Retrieved December 3, 2008 from `http://www.sseriga.edu.lv/library/working_papers/FT_2005_5.pdf`.

Bollen, K. A. (1989), *Structural Equations with Latent Variables*, John Wiley & Sons, New York, Chichester, Brisbane, Toronto, Singapore.

Bowles, S. (2008), 'Policies designed for self-interested citizens may undermine "the moral sentiments": Evidence from economic experiments', *Science* **320**(5883), 1605–1609. Retrieved on February 11, 2009 from `http://www.sciencemag.org/cgi/reprint/320/5883/1605.pdf`.

Bowles, S. & Hwang, S.-H. (2008), 'Social preferences and public economics: Mechanism design when social preferences depend on incentives', *Journal of Public Economics* **92**(8-9), 1811–1820. Special Issue: Happiness and Public Economics.

Brēķis, E. (2007), Latvia's Underground Economy Modelling With An Emphasis On Tax Policy, Summary of the promotion paper, University of Latvia, Riga. Retrieved on December 3, 2008 from `http://www3.acadlib.lv/grey/grey1.idq?CiRestriction=economy`.

Breusch, T. (2005), Estimating the underground economy using MIMIC models. Retrieved November 26, 2008 from http://129.3.20.41/eps/em/papers/0507/0507003.pdf.

Buehn, A. & Schneider, F. (2008), MIMIC models, cointegration and error correction: An application to the French shadow economy. IZA Discussion Paper No. 3306. Retrieved July 17, 2008 from http://ftp.iza.org/dp3306.pdf.

Chaudhuri, K., Schneider, F. & Chattopadhyay, S. (2006), 'The size and development of the shadow economy: An empirical investigation from states of India', *Journal of Development Economics* **80**, 428–443.

COM(2007) 628 final (24.10.2007), 'Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions: Stepping up the fight against undecleared work'. Retrieved on January 5, 2009 from http://ec.europa.eu/employment_social/news/2007/oct/undeclared_work_en.pdf.

Commission Decision No 2006/595/EC (2006), 'Commission Decision of 4 August 2006 drawing up the list of regions eligible for funding from the Structural Funds under the Convergence objective for the period 2007–2013', Official Journal of the European Union, L 243/44. Retrieved on August 20, 2009 from http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:243:0044:0046:EN:PDF.

Council Regulation (EC) No 1083/2006 (2006), 'Council regulation of 11 July 2006 laying down general provisions on the European Regional Development Fund, the European Social Fund and the Cohesion Fund and repealing regulation (EC) No 1260/1999', Official Journal of the European Union, L 210/25. Retrieved on August 20, 2009 from http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:210:0025:0078:EN:PDF.

Dell'Anno, R. (2007), 'The shadow economy in Portugal: An analysis with the MIMIC approach', *Journal of Applied Economics* **X**(2), 253–277.

Dell'Anno, R., Gómez-Antonio, M. & Alañon-Pardo, A. (2007), 'The shadow economy in three Mediterranean countries: France, Spain and Greece. A MIMIC approach', *Empirical Economics* **33**(1), 51–84.

Dell'Anno, R. & Schneider, F. (2003), 'The shadow economy of Italy and other OECD countries: What do we know?', *Journal of Public Finance and Public Choice* **XXI**(2-3), 97–120.

Dell'Anno, R. & Schneider, F. (2006), 'Estimating the underground economy by using MIMIC models: A response to T. Breusch's critique', Working Paper No. 0607, Department of Economics, Johannes Kepler University of Linz. Retrieved November 26, 2008 from `http://www.econ.jku.at/papers/2006/wp0607.pdf`.

Eilat, Y. & Zinnes, C. (2000), 'The evolution of the shadow economy in transition countries: Consequences for economic growth and donor assistance', CAER Discussion paper No. 83. Retrieved on July 10, 2008 from `http://pdf.usaid.gov/pdf_docs/PNACK691.pdf`.

Enste, D. H. (2008), Informelle Ökonomie: Umfang, Struktur und Ursachen, *in* D. Henckel, G. Spars & F. Wukovitsch, eds, 'Arbeiten in der Grauzone: Informelle Ökonomie in der Stadt', Peter Lang, Frankfurt am Main, chapter 1, pp. 23–45.

ESA (1995), 'European System of Accounts ESA 1995'. Retrieved on August 24, 2009 from `http://circa.europa.eu/irc/dsis/nfaccount/info/data/esa95/en/esa95en.htm`.

European Commission (2009), *VAT Rates Applied in the Member States of the European Community. DOC/2401/2009 - EN*, European Commission, Taxation and Customs Union, Brussels. Retrieved on August 28, 2009 from http://ec.europa.eu/taxation_customs/resources/documents/taxation/vat/how_vat_works/rates/vat_rates_en.pdf.

Eurostat (2006), *Taxation trends in the European Union: Data for the EU Member States and Norway. 2009 edition*, Research in official statistics, Office for Official Publications of the European Communities, Luxembourg. Retrieved on August 28, 2009 from http://ec.europa.eu/taxation_customs/resources/documents/taxation/gen_info/economic_analysis/tax_structures/2009/2009_full_text_en.pdf.

Feige, E. L. & Urban, I. (2008), 'Measuring underground (unobserved, non-observed, unrecorded) economies in transition countries: Can we trust GDP?', *Journal of Comparative Economics* **36**(2), 287–306.

Frey, B. S. & Weck, H. (1983), 'Estimating the shadow economy: A 'naive' approach', *Oxford Economic Papers* **35**(1), 23–44.

Frey, B. S. & Weck-Hanneman, H. (1984), 'The hidden economy as an 'unobserved' variable', *European Economic Review* **26**, 33–53.

Giles, D. (1997), The hidden economy and the tax-gap in New Zealand: A latent variable analysis. Retrieved on 09.10.2008 from http://web.uvic.ca/econ/wp9807.pdf.

Giles, D. (1999), 'Measuring the hidden economy: Implications for econometric modelling', *The Economic Journal* **109**(456), F370–F380.

Giles, D. E. & Tedds, L. M. (2002), *Taxes and the Canadian Underground Economy*, number 106 *in* 'Canadian Tax Paper', Canadian Tax Foundation, Toronto.

Grabowski, M. H. (2003), Informal labor market and informal economy during economic transition: The Polish perspective, *in* B. Belev, ed., 'The Informal Economy in the EU Accession Countries: Size, Scope, Trends and Challenges to the Process of EU Enlargement', Center for the Study of Democracy, Sofia, chapter 7, pp. 175–186.

Helberger, C. & Knepel, H. (1988), 'How big is the shadow economy? : A re-analysis of the unobserved-variable approach of B.S. Frey and H. Weck-Hannemann', *European Economic Review* **32**(4), 965–976.

ISTAT (2008), La misura delleconomia sommersa secondo le statistiche ufficiali. Instituto nazionale di statistica. Retrieved on April 9, 2009 from `http://www.istat.it/salastampa/comunicati/non_calendario/20080618_00/testointegrale20080618.pdf`.

Jöreskog, K. G. & Goldberger, A. S. (1975), 'Estimation of a model with multiple indicators and multiple causes of a single latent variable', *Journal of the American Statistical Association* **70**(351), 631–639.

Le Gallo, J. & Dall'erba, S. (2006), 'Evaluating the temporal and spatial heterogeneity of the European convergence process, 1980-1999', *Journal of Regional Science* **46**(269–288), 2.

LFMI (2005), 'The 15th survey of the Lithuanian economy 2004/2005 (2)', Lithuanian Free Market Institute, 03.05.2005. Retrived on December 4, 2008 from `http://www.freema.org/index.php/research/the_15th_survey_of_the_lithuanian_economy_20042005_2/2768`.

McMillen, D. P. (1996), 'One hundred fifty years of land values in Chicago: A nonparametric approach', *Journal of Urban Economics* **40**(1), 100 – 124.

Mróz, B. (2005), The shadow economy in Poland and its socio-economic implications, *in* S. Bukowski, ed., 'The Dilemmas of Economic Policy, Development and Regional Economic Integration', WSFiB, Poland, Radom, pp. 167–176. Retrieved on February 9, 2009 from `http://www.wydawnictwa.wsfib.edu.pl/Polska_w_UE/mr%F3z.pdf`.

Nastav, B. & Bojnec, Š. (2007), 'Shadow economy in Slovenia: The labour approach', *Managing Global Transitions* **5**(2), 193–208.

Schneider, F. (2003), The size and development of the shadow economies and shadow economy labor force of 22 transition and 21 OECD countries: What do we really know?, *in* B. Belev, ed., 'The Informal Economy in the EU Accession Countries: Size, Scope, Trends and Challenges to the Process of EU Enlargement', Center for the Study of Democracy, Sofia, chapter 2, pp. 23–61.

Schneider, F. (2007), 'Shadow economies and corruption all over the world: New estimates for 145 countries', *Economics: The Open-Access, Open-Assessment E-Journal* **1**(2007-9). Retrieved on December 4, 2008 from `http://www.economics-ejournal.org/economics/journalarticles/2007-9`.

Schneider, F. & Bajada, C. (2005), An international comparison of underground economic activity, *in* C. Bajada & F. Schneider, eds, 'Size, Causes and Consequences of the Underground Economy: An International Perspective', Ashgate, Aldershot and Burlington, chapter 5, pp. 73–106.

Schneider, F. & Enste, D. H. (2000), 'Shadow economies: Size, causes and consequencies', *Journal of Economic Literature* **38**(1), 77–114.

Thomas, J. J. (1999), 'Quantifying the black economy: 'measurement without theory' yet again?', *The Economic Journal* **109**(456), F381–F389.

United Nations (2008), *Non-observed Economy in National Accounts: Survey of Country Practices*, United Nations Economic Commission for Europe, New York and Geneva.

Williams, C. C. (2004), *Cash-in-Hand Work. The Underground Sector and The Hidden Economy of Favours*, Palgrave Macmillan, Basingstoke, Hampshire and New York.

Williams, C. C. & Windebank, J. (1998), *Informal Employment in the advanced economies: implications for work and welfare*, Routledge, London and New York.

Zellner, A. (1970), 'Estimation of regression relationships containing unobservable independent variables', *International Economic Review* **11**(3), 441–454.