The Dark, and Independent, Side of the Italian Labour Market

AUTHOR[•]

"(…) do not give him a fish;
take him to the water and
teach him to catch fish."
Confucius

Abstract

In many countries the shadow employment has a very high priority among policymakers. A new time series for this component of the labour market has been recently released by the Italian institute of statistics. Taken together they give the motivation and the occasion for a fresh analysis of the Italian labour market over the last two decades. The aim and the contribution of this paper is to highlight some stylised facts about the links between the two sides of the labour market, the dark and the regular. Results from "exhaustive" VAR/VEC models suggest that there are no connections (causal relationships, feedbacks, contemporaneous correlation) between these two time series. In this sense, we could correctly refer to the undeclared work as an "independent" side of the Italian labour market. I interpret these results as providing empirical support for the ineffectiveness of labour policies in converting black employment into regular one.

JEL classification: C53, E26, H26, J30.

Keywords: Underground economy; VAR models; Dual labour market.

[•] I would like to acknowledge the able assistance of Mrs. XXXX and helpful comments from an anonymous referee. All errors are solely those of author as are the opinion expressed herein.

1. Introduction

The non-observed sector of the economy has neither a commonly accepted definition, nor even a commonly used name. A plethora of terms (underground, undeclared, moonlight, hidden, irregular, shadow, black, etc.) have been used to call it. All of them are suggestive of a particular aspect of the phenomenon, which is manifold. I will indifferently use here some of these adjectives but as the findings of this paper suggest, in the Italian case another suitable label could be "independent". Since I use data drawn from the Italian national institute of statistics (Istat), the definition of black economy is the "official" one. Thus, the hidden sector which I refer to represents (U.N. *et al.*, 1993) the area of (legal) production activities that are not directly observed due to reasons of economic nature (deliberate desire to avoid taxes and/or to avoid observing the law provisions concerning the labour market) and/or statistical nature (*e.g.* due to the failure to fill out the administrative forms or statistics questionnaires).

There are several important reasons to analyse the potential links between regular and irregular activity. In a highly indebted system may be useful to ask oneself if fiscal policy can go on with a long sequence of surpluses, hoping that the regular sector does not sensitively react. A "mass escape" from the regular sector would dramatically reduce government revenues worsening the public budget situation. On the positive side, in a climate of economic stagnation and decline the underground economy may serve a useful economic and social function providing jobs to many of willing workers. Economic literature suggests that tax rates are negatively associated both with the labour supply and/or with the tax evasion (Schneider and Enste, 2000). The linkages can derive from labour market policies as well. In a paper by Boeri and Garibaldi (2002) it is argued that any unemployment reducing policy will endogenously reduce shadow employment, while it is very difficult to reduce shadow employment without increasing unemployment. By and large, effective policies aimed (not necessarily in a direct way) to convert hidden labour into regular one should increase the time series correlation between them. In this sense, the underground economy could be seen as a resource rather than a constraint from the policymakers' point of view. The tax amnesties implemented in Italy during the last decades are suggestive episodes as regard to this possibility.

Against this background, this paper tries to highlight some stylised facts about the links between the two sides of the Italian labour market, the dark and the regular. Due to limited data availability, very few works address these issues with a medium-term perspective (some exemptions are Bovi, 1999; Busato and Chiarini, 2004). The occasion to improve and to update the little available evidence is given by the recent Istat realise reporting annual data for the non-observed labour input (full-time equivalent[1], FTE, units) throughout the period 1980-2002 (Istat, 2004). Working with these new data, I examine the relationships between undeclared and regular FTE units via a time series analysis. Missing a consolidated economic theory and due to small sample issues, I chose to be as agnostic as possible. In other words, I do not carry out preliminary univariate analyses of the two (short) annual time series. Instead, I perform unit-root tests on the total FTE, available at quarterly frequency, to lessen the set of bivariate vector autoregression (VAR) models which should include the correct specification. Then, I perform innovation accounting and impulse response functions analyses in order to see if, how and how much the two faces of the market interact. Somewhat puzzling (Busato and Chiarini, 2004; Lucifora, 2003; Signorelli, 1997), all models point to the same conclusion – the two sectors are orthogonal or, as suggested in the title, independent. No Granger causality, no contemporaneous correlation, and no shocks transfer from one market to another emerge from the data. This evidence is congruent with an interesting normative interpretation. During the last twenty years, policies had not a significant influence in transforming hidden employment into regular one.

The paper is organised as follow. The next section deals with the Istat method and the data issues, section 3 presents preliminary univariate tests, and section 4 is concerned with VARs analyses. Concluding remarks are collected in the final section.

[1] The number of the full time equivalent units are equal to the number of jobs corresponding to full time. The total of full time equivalent units is obtained by the sum of (primary and secondary) full-time jobs and part-time jobs transformed into full-time units (Eurostat, 1996).

2. The Non-Observed Italian Economy. The Labour Input Method and its Drawbacks.

The source of the data is Istat, thus the definition of the non-observed sector is the "official" one. The hidden production represents, according to the System of National Accounts (SNA) definition (U.N. *et al.*, 1993), the area of (legal) production activities that are not directly observed due to reasons of economic nature (deliberate desire to avoid taxes and/or to avoid observing the law provisions concerning the labour market) and/or statistical nature (*e.g.* due to the failure to fill out the administrative forms or statistics questionnaires). Istat claims that non-observed does not means non-measured (Calzaroni, 2000; Baldassarini and Pascarella, 2003), and its estimation approach is known as the labour input method (OECD, 2002). Briefly, it consists in i) the use of sources and survey techniques that make possible to measure the weight of unregistered work (this is achieved primarily by using labour status particulars declared by respondents in the household surveys: it is assumed that individuals have less reasons than enterprises to conceal the nature of their work); ii) the correction of the under-reporting of income by the enterprises through adjustments of the per capita production and value added values declared by the small production units (fewer than 20 employees) and iii) the checks for the consistency of the economic aggregates through the balancing of the resources and uses made at the level of each industry. As a result, Istat publishes annual estimates of the irregular input of labour and GDP. The main focus of this paper is the labour market, therefore I deal with labour input (full time equivalent units) data. On the other hand, this time series is the only one available for the whole (1980-2002) sample.

While it is simple to describe, the practical application of the labour input method is more difficult (OECD, 2004):

- Labour force surveys provide estimates of the numbers of workers, while data from enterprise surveys usually refer to the number of jobs. The two sets of data must therefore be standardised by converting them to comparable units such as hours worked or full-time equivalents.

- The comparison between the two independent sets of data needs to be made at as detailed a level as possible. Ideally this should be by size-classes as well as by detailed kinds of activities.

- The method depends crucially on the availability of comprehensive estimates of labour inputs. Istat considers that the results of the household labour force survey, supplemented by demographic data, provide such estimates[2]. The Italian survey collects information on the kind of activity, hours worked, and the approximate number of employees in the enterprises where the respondents work. The survey also collects information on secondary jobs, which are relatively common in Italy for persons whose main job is with the government. In addition, Istat tries to take into account non-resident undocumented foreigner workers, which can not be observed directly by the usual sources used to uncover other kinds of black economy.

Needless to say, although the method is internationally recognised to be a very good one (OECD, 2002) it is not immune from concerns and problems. Even if it is reasonable to assume that individuals have less reasons than enterprises to conceal the nature of their work, Boeri and Garibaldi (2002) point out that if employees cooperate in shadow activities they may decide not to declare to be working. As reported in their paper, a joint Istat-Fondazione Curella survey reports that about 25% of the black economy is wrongly assigned to the inactive status by the labour force survey. Also, some individuals who indicate to their interviewer that they are self-employed may actually be labouring in the underground economy. A study of the US internal revenue service[3] found that 47% of the workers who were classified as independent contractors did not report any taxable income. Then, one can speculate that unemployment could be overstated in the surveys because respondents who should have been classified "out of the labour force" are fearful that they would lose benefits unless they indicated they were looking for work. On the other hand, some writers on the underground economy have pointed out that the low labour force participation might

[2] Even if Istat knows (and surveys) only regular firms, from households' answers it can detect irregular workers engaged both for regular and for irregular firms.

[3] Budget of the United States Government, 1984, p. 5–120.

reflect concealment of some employment activities (*e.g.* Gutmann, 1979). Furthermore, as Tanzi suggests (Tanzi, 1981), the first issue for the irregular sector worker when approached by the interviewer, is whether to become a respondent. It seems reasonable to assume that he is more likely to be a non-respondent than he would be if he were not in the irregular sector. In the Istat approach, non-respondent are included in the "statistical underground", which is allocated to the observed economy. In 1998 the percentage of non response was 3% of total GDP (Istat, 1998). On the other hand, the respondent may want to avoid telling anyone the truth about sources of income, and so will have concocted a convenient story intended to arouse the least suspicion. A non-specific but legitimate sounding job would appear the easiest way out for those individuals. In this case, the family survey data are larger than the firms'ones.

Evidence reported in Di Nardo *et al.*, (2000) may give an idea of the potential bias. In this paper is described a survey carried out in San Giuseppe Vesuviano, a town near Naples known for its widespread black employment. The standard method was that of the census survey, but conducted three times for the same universe in a period of a few months. Once by an interviewer not known to the local area; once by an interviewer who was known, using the indirect method of contacting 'key observers'; and once by a known interviewer using the direct, door-to-door survey method, but establishing trust with the respondents and exploiting (fortuitously) her particular personal situation - she had to finish her thesis and she was pregnant. Where the standard method produced a result of 31.1% and that of the "informed persons" one of 35.8%, the third approach got 43.7%. While this kind of research may suggest that the bias could be significant, its scientific content and replicability is questionable. Just to mention, what about other "particular" personal situations? Then, it is really hard to imagine how this method could be structurally implemented in the system of national accounts.

Having said that, let us turn the attention to the data. The two time series, non-observed labour input (FTE_black) and regular labour input (FTE_reg), are displayed in Appendix 1. Figure 1 shows that the shadow employment has been raising more quickly in the first decade (for a total increase of about 24%), than thereafter (for a total increase of about 12%, excluding the 2002). Even recently, the problem of a booming underground sector is often evoked in political circles. On the contrary, data seem to suggest that the real question now is the level of the phenomenon, not its growth. To the extent the shadow workers have a lower productivity than the regular ones, these dynamics are consistent with the reversion of the jobless economic growth started in the mid-1990s (Bertola and Garibaldi, 2003). Up to the 1998, the behaviour of the irregularity ratio (black employment on total FTE) is more uniform (Figure 2). As refer to FTE_reg, it can be noticed the effect of the 1993 crisis and the strong recovery starting in 1998. The irregularity ratio clearly show the impact of the legalisation-regularisation of about 700,000 illegal immigrants in 2002 (Istat, 2004).

3. Univariate analysis

One common way to properly analyse the empirical relationships between time series is the vector autoregression approach. The first necessary step before validly estimating and using a VAR model is the analysis of the stochastic properties of the involved series. The attention devoted to this topic is well deserved for several reasons. On the one hand in contrast to stationary or trend stationary time series, models with a stochastic trend are persistent in the sense that shocks have permanent effects on the values of the process; otherwise stated, they have time dependent variances that go to infinity with time. Second, when a series is used in regressions with other variables the interpretation of the regression results can depend on whether the variables involved are trend (TS) or difference (DS) stationary. This phenomenon is related to the "nonsense" and "spurious" regression literature (Phillips, 1986) due to Yule (1926) and Granger and Newbold (1974).

It is also well known that unit root tests are based on asymptotic critical values. One expects in finite samples that the use of asymptotic critical values will result in over-rejection, and twenty-three (1980-2002) observations are definitively a finite sample. I address this potential problem with a two-steps procedure. First, I test if the total labour input time series (FTE_tot=FTE_black+FTE_reg) is an integrated process. I can be sufficiently sure about the robustness of this result because FTE_tot is disposable at quarterly frequency, thus I have about ninety observations. In this first step I pin down the stochastic properties of the total labour input. Then, I take advantage of the algebra of integrated processes (Granger and Hallmann, 1991) to shrink the "exhaustive" VAR analysis only to the remaining (plausible) models. This latter analysis is the second step of the procedure (section 4).

I perform two unit root tests. The first (NP) was worked out by Ng and Perron (2001) and it is four-fold. It yields both substantial power gains and a lower size distortions over the standard unit root tests, maintaining the null of unit root. NP offer four test statistics based on the Generalized Least Squared (GLS) detrended data y_t^d. Altogether these statistics are enhanced versions of Phillips-Perron Z_α and Z_t statistics (1988), the Bargava R_1 statistic (1986), and the Elliot *et al.* Point Optimal statistic (1996):

$$MZ_\alpha = \left(T^{-1}(y_t^d)^2 - f_0\right)/2\kappa$$

$$MSB = \left(\kappa/f_0\right)^{1/2}$$

$$MZ_t = MZ_\alpha \; X \; MSB$$

$$MPT = \bar{c}^2 \kappa - \bar{c}T^{-1}(y_t^d)^2)/f_0 \qquad \text{(if exogenous = constant)}$$

$$MPT = \bar{c}^2 \kappa + (1-\bar{c})T^{-1}(y_t^d)^2)/f_0 \qquad \text{(if exogenous = constant, trend)}$$

where $\kappa = \sum_{t=2}^{T} (y_{t-1}^d)^2)/T^2$ and f_0 is an estimate of the residual spectral density at the zero frequency[4]. The choice of the autoregressive truncation lag, p, is critical for correct calculation of f_0. Here p is chosen using the modified AIC suggested by Ng and Perron (2001).

The second is the KPSS test (Kwiatkowski, *et al.* (1992)), which can be thought as complementing the NP one because it tests the null hypothesis that FTE_tot is a TS stochastic process. Suppose the NP test fails to reject the unit root null because of low power. The KPSS test which has (trend) stationarity as the null should indicate the data have no unit roots. On the other hand, if the KPSS test rejects the trend stationarity null,

[4] The frequency zero spectrum method used is the AR-GLS detrended.

then we have stronger evidence for unit root persistence. That is, consistent results from NP and KPSS tests yield more persuasive evidence on data persistence, while conflicting results indicate[5] uncertainty associated with the interpretation of the individual test outcomes. The KPSS test is based upon the residuals from the OLS regression of y_t on the exogenous variables x_t:

$$y_t = x_t'\delta + u_t$$

The LM statistic is be defined as:

$$LM = \sum_t^n S(t)^2 / (T^2 f_0),$$

where f_0 is an estimator[6] of the residual spectrum at frequency zero and where S(t) is a cumulative residual function:

$$S(t) = \sum_{r=1}^{t} \hat{u}_r$$

based on the residuals $\hat{u} = yt - x'\hat{\delta}(0)$.

Table 1. Unit root tests on (log-)labour input (FTE units, quarterly data 1980:1-2002:4)

		MZa		MZt		MSB		MPT		KPSS*	
Test Statistics	FTE_tot**	-3.34		-1.22		0.36		25.82		0.449	
	D(FTE_tot)* **	**-9.41**		**-2.14**		**0.23**		**2.73**		**0.190**	
	Confidence	**	***	**	***	**	***	**	***	**	***
Critical Values	1%	-23.8	**-13.8**	-3.42	**-2.58**	0.143	**0.174**	4.03	**1.78**	0.216	**0.734**
	5%	-17.3	**-8.1**	-2.91	**-1.98**	0.168	**0.233**	5.48	**3.17**	0.146	**0.463**
	10%	-14.2	**-5.7**	-2.62	**-1.62**	0.185	**0.275**	6.67	**4.45**	0.119	**0.347**

Lag length criterion Modified AIC; $D(x)=x-x_{-1}$; *H0: TS process; **constant and trend included; ***only constant.

Table 1 indicates that the (log) level of the labour input time series is a DS process. NP M-tests fail to reject the null of a unit root, while KPSS strongly rejects the null of stationariety[7]. On the other hand, the first difference of FTE_tot is stationary.

This outcome implies that FTE_black and FTE_reg can be neither both stationary nor cointegrated with a cointegration coefficient equal to one. In fact, Granger and Hallmann (1991) show that for a pair of independent variables holds (omitting coefficients and using a widespread notation):

$$I(0) + I(0) = I(0);\ I(1) + I(0) = I(1);\ I(1) + I(1) = I(1).$$

If the two series are cointegrated, then I(d)+ I(d) = I(d-1), where d is the order of integration.

So, we remain with four possible VAR models linking the two components of FTE_tot[8] (in log, $D(x)=x-x_{-1}$):

[5] Conflicting results might also indicate the presence of fractionally integrated processes.
[6] The frequency zero spectrum method used is the Kernel-Bartlett sum-of-covariances.
[7] Standard ADF and nonparametric Phillips-Perron tests confirm the I(1) nature of FTE_tot.
[8] A unit root in FTE_tot could be validly consistent with the cointegrated and I(2) nature of its two parts. I can not rule out this event because the labour input is a stock variable, thus its first difference may be an I(1) flow. As somewhat expected, tentative unit root tests show mixed evidence on the I(0) vs I(1) nature of FTE_black and FTE_reg.

Model 1: FTE_black; D(FTE_reg);

Model 2: D(FTE_black); FTE_reg;

Model 3: D(FTE_black); D(FTE_reg);

Model 4: FTE_black; FTE_reg (VECM).

The first two bivariate VARs are consistent with the case “only one FTE component is I(1)”, Model 3 would be valid if both FTE_black and FTE_reg were DS but non cointegrated, the last model account for the possibility of a cointegrating relationship between them. Altogether, these models make up an “exhaustive” VAR analysis including the “true” model. To uncover robust results while operating under this “veil of ignorance and caution”, it needs to obtain “all-models” information. Otherwise stated, since I do not know which model is correctly parameterised in order to draw any conclusion evidence pointing in one direction independently of the model is necessary. On the positive side, uniform findings turn out to be very robust.

4. Vector Autoregression Analysis

Through the analysis of the covariances, the VAR approach allows us to see if one market has a tendency to lead the other, if there are feedbacks between them, and how do impulses (shocks, innovations) transfer from one sector to another. The VAR approach (Sims, 1980) sidesteps the need for structural modelling by treating every endogenous variables in the system as a function of the lagged values of all the endogenous variables in the system. Consider the VAR(p) model

$$\Phi(L)\boldsymbol{y}_t = \varepsilon_t$$

where $\Phi(L) = \boldsymbol{I} - \Phi_1 L - \Phi_2 L^2 - \ldots - \Phi_p L^p$.

The innovations ε_t may be contemporaneously correlated, but are uncorrelated with their own lagged values and uncorrelated with all of the right-hand side variables. Since only lagged values of the endogenous variables appear on the right-hand side of the equations, simultaneity is not an issue and OLS yields consistent estimates. Moreover, even though the innovations may be contemporaneously correlated, OLS is efficient and equivalent to GLS since all equations have identical regressors. A basic assumption in the above model is that the residual vector follows a multivariate white noise. Also, in order that the VAR-model is stationary, it is required that roots of $| \boldsymbol{I} - \Phi_1 z - \Phi_2 z^2 - \ldots - \Phi_p z^p | = 0$ lie outside the unit circle. Provided that the stationary conditions hold we have the vector moving average representation of $\boldsymbol{y}_t$ as

$$\boldsymbol{y}_t = \Phi^{-1}(L)\varepsilon_t = \varepsilon_t + \sum_{i=1}^{\infty} \psi_i \, \varepsilon_{t-i}$$

where ψ_i is an m×m coefficient matrix. The ε_t's represent shocks in the system. Suppose we have a unit change in ε_t then its effect in $\boldsymbol{y}$ s periods ahead is $\frac{\delta y_{t+s}}{\delta \varepsilon_t} = \psi_s$.

Accordingly the interpretation of the ψ matrices is that they represent marginal effects, or dynamic multipliers, or the model's response to a unit shock (or innovation) at time point t in each of the variables. The response of y_i to a unit shock in y_j is given by the sequence, known as the impulse response function (IRF),

$$\psi_{ij,1}, \psi_{ij,2}, \psi_{ij,3}, \ldots,$$

where $\psi_{ij,k}$ is the ijth element of the matrix ψ_k (i, j = 1, . . . , m). Generally speaking an IRF traces the effect of a one-time shock to one of the innovations on current and future values of the endogenous variables. Otherwise stated, the IRFs trace out how the variables will deviate from the path predicted by the model if there is a forecast error with respect to a specific equation at time t. Unforeseen movements in y_j are referred to as shocks and the state of the economy at the time t+m as responses. However, unless the error covariance matrix $E(\varepsilon_t \varepsilon_t')$ is a diagonal matrix, the shocks will not occur independent from each other. The conventional practice in the VAR literature is to single out the individual effects by first orthogonalise the error covariance matrix, *e.g.* by Cholesky decomposition, such that the new residuals become contemporaneously uncorrelated with unit variances. Unfortunately orthogonalization is not unique in the sense that changing the order of variables in $\boldsymbol{y}$ changes the results. The economic theory may be used to solve the ordering issue. As usual the approach I follow here is agnostic, and it is based on examining the two possible orderings (because of the bivariate VAR) to see whether the resulting interpretations are consistent. Since in a bivariate model the Granger-causality implies that one variable must react to a shock of the other, within this framework I can address the causality issues[9] as well.

The uncorrelatedness of the new residuals allows the error variance of the s step-ahead forecast of y_{it} to be decomposed into components accounted for by these shocks. Because the innovations have unit variances, the components of this error variance accounted for by innovations to y_j is given by

$$\sum_{k=0}^{s} \psi^{*\,2}_{ij,k}$$

where ψ^*_i is the orthogonalised version of ψ_i. Comparing this to the sum of innovation responses we get a relative measure how important variable y_j innovations are in the explaining the variation in variable i at different step-ahead forecasts, *i.e.*,

$$R^2_{ij,s} = 100 \frac{\sum_{k=0}^{s-1} \psi^{*\,2}_{ij,k}}{\sum_{h=1}^{m} \sum_{k=0}^{s-1} \psi^{*\,2}_{ih,k}}$$

Thus, while IRFs trace the effects of a shock to one endogenous variable on to the other variables in the VAR, variance decomposition separates the variation in an endogenous variable into the component shocks to the VAR. Clearly, even the variance decomposition results depend on the ordering when there is contemporaneous correlation between the residuals. Again, for the robustness of the findings I replicate the two possible orderings of the bivariate VAR.

Another useful and workable set of experiments within the present statistical-atheoretical context is the analysis of the generalised IRFs. Pesaran and Shin (1998) have suggested a theoretically neutral way of deriving impulse responses that takes into account the information on the correlation of errors contained in the error covariance matrix. These authors construct an orthogonal set of innovations that does not depend on the VAR ordering. The generalized impulse responses from an innovation to the j^{th} variable are derived by applying a variable specific Cholesky factor computed with the j^{th} variable at the top of the Cholesky ordering. It should be noticed that the generalised response profiles derived in this way are not conveying information about economic causation among the variables. The exercise can be thought of as tracing out how the observation of a forecast error in one equation of the system would lead to revisions in the forecast path of all model variables.

Finally, in the absence of contemporaneous exogenous variables the disturbance variance–covariance matrix contains all the information about contemporaneous correlations among the variables. Thus, the more the results are independent from the ordering, the more the innovations are contemporaneously uncorrelated. The IRFs and the (low) residuals contemporaneous correlation reported in Appendix 3 give a visual idea of the effect of the latter on the former. In that, we have another useful indication from this analysis.

Estimating the parameters in a VAR requires that the involved variables be covariance stationary, meaning that their first two moments exist and are time invariant. If they are not covariance stationary, but their first-differences are (see section 3), a vector error-correction model (VECM) can be used to perform the same set of experiments. In fact, a VECM is nothing else than a restricted VAR designed for use with nonstationary

[9] As well known, the Granger causality does not imply that a variable x is the effect or the result of a variable y. It just measures precedence and information content.

series that are known to be cointegrated. Basically, unlike unrestricted VARs, the VECM has cointegration relations (equations) built into the specification so that it restricts the long-run behaviour of the endogenous variables to converge to their cointegrating relationships while allowing for short-run adjustment dynamics. The cointegration term is known as the error correction term since the deviation from long-run equilibrium is corrected gradually through a series of partial short-run adjustments. Otherwise equal VECMs may differ in their deterministic part. Actually, the series may have nonzero means and deterministic trends as well as stochastic trends. Similarly, the cointegrating equations may have intercepts and deterministic trends. Since FTE_black and FTE_reg seem to be trending series (see Appendix 1) and for robustness, I perform two plausible kinds of VECMs (Models 4 and 5 in Appendix 2). In the first VECM, the cointegrating equation[10] has only the intercept; in the second, the cointegrating equation has both the intercept and the linear trend. Unlike the unrestricted VAR case, the innovation response plots for VECMs are not endowed with standard error bands since some of the maximum likelihood parameter estimates have nonstandard asymptotic distributions.

Despite just one model is correctly specified, the analyses of the VAR residuals reported in Appendix 2 (tables 1a-5a) suggest that all the models provide a fair description of the information in the data. After controlling for one or two outliers each and every VAR satisfies residual assumptions, while there is no sign of instability[11]. The limited sample available is likely to play a role, suggesting that it is worth performing several models. The following figures (Appendix 3) plot the relative mean estimates of the (Cholesky and Generalised) impulse response functions and show the variance decomposition outcomes. The pure shape of impulse functions is not fully informative of whether a detected reaction path is also meaningful in a statistical sense. Thus, I also display the upper and lower limits of a 95% Monte Carlo band[12]. Clearly, if these bands contain the zero line one can conclude that there is evidence of no reaction. All the unrestricted VARs have the same exogenous variable, namely a constant. Not reported sensitivity analyses conduct adding a linear or a quadratic trend do not substantially change the stylised facts that emerge. These latter are reported and commented in the following section.

[10] Obviously, within a bivariate VAR framework at most one cointegrating relation is possible.
[11] All the inverse roots of the characteristic AR polynomial lie inside the unit circle (Appendix 2).
[12] As already mentioned, for VECMs it is not possible to calculate standard errors.

5. Concluding Remarks

The aim of this paper is to study if, how, and how much the dark and the regular side of the Italian labour market are linked. In doing that I exploit the recently released time series data for these two components, which are analysed in a VAR framework. Because of small sample issues (the annual data cover the period 1980-2002) it is hard to infer the correct location of unit roots, if any, in the system. That is why I "over-model", *i.e.* I estimate all plausible VAR specifications, which include the correct one. Then, I focus only on evidence pointing in one direction independently of the model. These robust results may be summarised in the following statements:

- the Italian labour market seems to be made up by two independent components, one regular, one irregular. In particular,
- unequivocal evidence shows that there is neither contemporaneous correlation, nor Granger causality between them. Also, shocks hitting one labour input do not lead to any reaction in the other.

It is worth recalling that statistical experiments have stronger ability in negating than in supporting the occurrence of an event. Thus, the outcomes are robust even in this respect.

The picture emerging from the empirical analysis is coherent with an interpretation particularly interesting for its normative content. Good labour market policies, directly[13] or indirectly targeted to reduce shadow employment, should convert black workers into regular ones (Comitato per l'emersione del lavoro non regolare, 2004). To be effective, they should therefore induce a negative correlation between these two kinds of input. Since data tell a different story, it turns out that despite several efforts throughout the last twenty years policies were not able to impinge on the two faces of the labour market in a significant way. This is somewhat consistent with the studies by Censis (2003), Bàculo (2002), Mateman and Renooy (2001), and with the widespread dissatisfaction expressed in the political circles over the situation that has been reached in the recent years[14].

A possible explanation is the following. Basically, the mainstream literature (Allingham and Sandmo, 1972 and followers) suggests that the hidden activity depends on its relative costs (the expected penalty) and revenues. The former depends on "institutional" factors (corruption, the efficiency/credibility of the public sector, etc.), the latter on "economic" determinants (such as the tax wedge). In Italy, shadow reducing policies have usually been aimed to hamper the revenues for black agents – often via tax amnesties/reductions. Less has been done to rise the cost to be underground. Sometimes these latter measures just consist in claiming "zero tolerance", which is in sharp contrast with the reiterate use of amnesties[15]. This raises a problem of credibility, whose importance in fighting the underground sector is documented in several papers. Johnson *et al.,* (1997) Friedman, *et al.* (2000), Kaufman (2003), and Bovi (2004) report cross-country evidence showing the pivotal role of institutions as a determinant of the "quit option". Bovi and Castellucci (2001) find similar findings for the Italian regions, while Boeri and Garibaldi (2002) argue that a total repression of the informal sector is not a credible threat because it is very difficult to reduce shadow employment without increasing unemployment. It is not (only) a carrot and stick matter, an investment in the overall institutional credibility/efficiency could be very productive even in this field. Also, peculiar factors may be at work behind the aggregate results and it is likely that generic policy efforts does not target all the causes of the undeclared employment[16]. The underground activity is a very a complex

[13] An example of these policies are the wage alignment agreements (contratti di riallineamento), implemented throughout the 90s. For a survey of direct and indirect policies implemented in Italy to combat underground labour, see Lucifora (2003).

[14] Starting from 1998 Italian governments seem to have increased their efforts and to have changed approach in tackling the problem Mateman and Renooy (2001). For instance, a "Committee for the Exposure of Undeclared Work" (Comitato per l'emersione del lavoro non regolare) was appointed in that year to discourage the black employment and to encourage the regular one.

[15] In passing, a survey by the World Bank shows that intermittent amnesties have had negative public revenue effects in all empirical studies that examine them (see: http://www1.worldbank.org/publicsector/tax/amnesties.html#how).

[16] *Inter alia*, if shadow activity is performed within a "network" (*i.e.* if firms buy and sell from other underground units), they can not emerge even if it is profitable from the individual point of view. In this case the policy target should be the whole network.

phenomenon (Brunetta and Ceci, 1998; Ministero del Tesoro, 2001; Meldolesi, 2003), and if a "one-policy-fits-all" approach may be ineffective, a step-by-step "fine tailoring" procedure may generate only a modest impact at the aggregate level, even in the medium term. This normative framework and empirical evidence shed some light on the suggested (Censis, 2003) hypothesis that firms respond more to tax credits for hiring new employees than to tax amnesties, in both cases without significantly abating the shadow employment. Then, it might be that the proposed policies did not still get profitable for firms to regularise black workers because their productivity is lower[17] (ISAE, 2002). To change their status they should structurally have a lower wage. As suggested (Boeri, 2002), wage policies could help in this dual market.

Problems may also stem from the labour supply-side (Ministero del Lavoro, 1987). Undertaking programs and moving in general towards the reduction of revenues for black agents may be hampered by the presence of means-tested benefits and by the prohibition for civil servants and retirees to cumulate other labour incomes. This latter condition just leaves no choice - black or nothing, in that preventing links between the two sides of the labour market. As for the former, the case of "social useful workers" and/or of beneficiaries of CIGS (extraordinary wages guarantee fund) working as irregulars, is instructive. They do not accept regularisation because they would loose the benefit. In addition, it is useful to spend some words about illegal immigrants. The only policy measure successfully transforming black jobs into regular ones seems to be the last legalisation[18] of illegal immigrants (Istat, 2004). A visual inspection of Figure 2 (Appendix 1) is suggestive of its impact despite the still marginal role of this item (about 16% on total FTE_black). A simple reason for that is that it is much more pressing to be legalised (*i.e.* to get the residency permit) than to be regularised (*i.e.* to work regularly). Moreover, it may be that the "institutional constraint" is more binding on illegal immigrants than on underground natives. A survey by Isae (2002) shows that to combat hidden employment in the personal services sector, featured by a widespread share of irregular foreigner workers, Italian households prefer a tax reduction to anti-crime intervention programs.

To conclude, some *caveat*. Results are data dependent, that is I can not exclude that the outcomes are biased because of measurement errors, such as those stemming from black workers with wrongly assigned labour status, or from non-exhaustive surveys. Then, I focus on full time equivalent units. While they measure the amount of labour input used in producing GDP, they do not allow to study the impact of policies on the number of jobs/persons. Finally, results are sample dependent and it could be just too early to judge the recently implemented policies. In longer samples these latter could induce the wanted negative correlation. Let us wait and see.

[17] See Busato and Chiarini (2004) for a two-sector RBC theoretical model with lower-productive black workers.
[18] In 2002, about 274,000 FTE units were regularised. Other legalisations (1990, 1996, 1998) are not so visible in the data (see Figure 2).

References

Allingham, M.G. and Sandmo A. (1972), "Income Tax Evasion: A Theoretical Analysis," in *Journal of Public Economics*, **I**: 323-338.

Bàculo L. (2002) *Valutazione dei contratti di riallineamento in due provincie meridionali*, in Isfol, La valutazione delle politiche del lavoro, Roma, F. Angeli.

Baldassarini A. and Pascarella C. (2003) The Italian approach to estimating the extent of the Non-Observed Economy: methods and results. Conference on The Non-Observed Economy: Measurement and Policy Issues Roma, 20-21 January. http://www.economia.uniroma2.it/noe/agenda_eng/base.htm

Bertola G. and Garibaldi P. (2003) "The Structure and History of Italian Unemployment", *CESifo working paper,* 907.

Bhargava A. (1986) "On the Theory of Testing for Unit Roots in Observed Time Series," *Review of Economic Studies*, **53**: 369-384.

Boeri, T. (2002) "Increasing the Size of the European labour Force: the Relevant Trade-offs," *Economic Survey of Europe*, UNECE, **2**.

Boeri T. and Garibaldi P. (2002) "Shadow Activity and Unemployment in a Depressed Labour Market," *CEPR Discussion Paper*, 3433.

Bovi M. (1999) "Un miglioramento del metodo di Tanzi per la stima dell'economia sommersa in Italia," *Rivista di Statistica Ufficiale dell'ISTAT*, **2**: 5-51.

Bovi M. (2004) "The (Underground) Wealth of Nations," *Politica Economica,* **1**: 117-139.

Bovi M. and Castellucci L. (2001) "Cosa sappiamo dell'economia sommersa in Italia al di là dei luoghi comuni? Alcune proposizioni empiricamente fondate," *Economia Pubblica*, **6** (English version: "What Do We Know About the Size of the Underground Economy in Italy Beyond the 'Common Wisdom'? Some Empirically Tested Propositions." *Quaderni CEIS*. Università di Roma Tor Vergata. **120**: 1-53).

Brunetta, R. and Ceci, A. (1998) "Il lavoro sommerso in Italia: cause, dimensioni e costi-benefici dell'emersione," *Economia Italiana*, Banca di Roma, **2**: 365-400.

Busato F. and Chiarini B. (2004) "Market and Underground Activities in a Two-Sector Dynamic Equilibrium Model," *Economic Theory*, **23**: 831-861.

Calzaroni M. (2000) "The exhaustiveness of production estimates: new concepts and methodologies," Paper presented at the ICES II (Second International Conference on the establishment surveys), Buffalo, 17-21 June.

Censis (2003) *Promuovere regolarità e trasparenze nel mercato del lavoro*, Dicembre, Roma

Comitato per l'emersione del lavoro non regolare (2004) *XI Relazione*, Luglio, Roma.

Di Nardo T. Caldarelli R Izzo M. (2000) "Le diverse verità della statistica: il caso delle forze di lavoro", *Economia Marche*, **2.**

Elliott G. Rothenberg T.J. and Stock J.H. (1996) "Efficient Tests for an Autoregressive Unit Root," *Econometrica* **64**: 813-836.

Eurostat (1996) *European System of National Accounts – ESA 1995*, Luxembourg, June.

Friedman E. Johnson S. Kaufmann D. and Zoido-Lobaton P. (2000) "Dodging the Grabbing Hand: The Determinants of Unofficial Activity in 69 Countries," *Journal of Public Economics*, **LXXVI**: 459-494.

Granger C.W.J. Hallmann J. (1991) "Nonlinear transformations of integrated time series," *Journal of Time Series Analysis* **12**: 207-224.

Granger C.W.J. and Newbold P. (1974) "Spurious Regressions in Econometrics," *Journal of Econometrics*, **2**: 111-120.

Gutmann P.M. (1979) “Are the Unemployed, Unemployed?” *Financial Analysts Journal*, 26-29.

Isae (2002) *Rapporto Trimestrale. Priorità nazionali: trasparenza, flessibilità, opportunità.* Aprile, Isae, Roma.

Istat (1998) “L’economia sommersa, problemi di misura e possibili effetti sulla finanza pubblica,” Hearing of the President of Istat. V° Commission. Chambers of Deputies. Roma.

Istat (2004) “La misura dell’occupazione non regolare nelle stime di contabilità nazionale: un’analisi a livello nazionale, regionale e retrospettiva a partire dal 1980,” *Comunicato fuori calendario*, 29th December, Istat, Roma.

Johnson S. Kaufmann D. and Schleifer A. (1997) “The Unofficial Economy in Transition,” *Brooking Papers on Economic Activity,* **II**: 159-239.

Kaufman D. (2003) “State capture and trends of the Non-Observed Economy in relations between state and the market,” Paper presented at the Conferenza internazionale Economia sommersa: misurazione e politiche Università di Tor Vergata, Roma, 20-21 January. http://www.economia.uniroma2.it/noe/agenda_eng/base.htm

Kwiatkowski D. Phillips P.C.B. Schmidt P. and Shin Y. (1992) “Testing the Null Hypothesis of Stationary against the Alternative of a Unit Root,” *Journal of Econometrics,* **54**: 159-178.

Lucifora, C. (2003) *Economia Sommersa e Lavoro Nero*, Bologna, il Mulino.

Mateman S. and Renooy P.H. (2001) *Undeclared Labour in Europe.Towards an integrated approach of combatting undeclared labour,* Final Report, Regioplan Publication.

Meldolesi L. (2003) “Policy for the Regulation of the Underground Economy and Employment,” *Review of Economic Conditions in Italy*, **I,** Capitalia, Roma.

Ministero del Lavoro (1987) *Report ’87 Labour and Employment Policies in Italy*, Roma.

Ministero del Tesoro (2001) *Linee guida per l’analisi degli effetti sull’occupazione degli interventi*, Dipartimento per le Politiche di Sviluppo e Coesione, Roma.

Ng S. and Perron P. (2001) “Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power,” *Econometrica*, **69(6)**: 1519-1554.

OECD (2002) *Measuring the Non-Observed Economy: A Handbook*, Paris, OECD.

OECD (2004) *Employment Outlook*, Paris, OECD.

Phillips, P.C.B. (1986) “Understanding Spurious Regression in Econometrics,” *Journal of Econometrics*, **33(31)**: l-40.

Phillips P.C.B. and Perron P. (1988) “Testing for a Unit Root in Time Series Regression,” *Biometrika*, **75**: 335–346.

Pesaran M.H. and Shin Y. (1998) “Impulse Response Analysis in Linear Multivariate Models,” *Economics Letters*, **58**: 17-29.

Signorelli M. (1997) “Uncertainty, Flexibility Gap and Labour Demand in the Italian Economy,” *Labour* **11(1)**: 141-175.

Sims C. (1980) “Macroeconomics and Reality,” *Econometrica*, **48**: 1-48.

Schneider F. and Enste D.H. (2000) “Shadow Economies: Size, Causes and Consequences,” *Journal of Economic Literature*. **38**: 77-114.

Tanzi, V. (1981) *The Underground Economy in the United States and Abroad*, Lexington, Heath and Co, MA.

U.N. Eurostat IMF OECD and Word Bank (1993) *System of National Accounts*, Brussels, Luxembourg, New York, Paris, Washington D.C.

Yule G.U. (1926) “Why do we sometimes find nonsense correlation between time series?” *Journal of the Royal Statistical Society*, **89**: 1-69.

APPENDIX 1. Dark and Regular Full Time Equivalent (FTE) Units

Figure 1. Dark (---FTE_black, right) and Regular (- - -FTE_reg, left) Labour Input Italy 1980-2002. (thousands of FTE units; source: Istat).

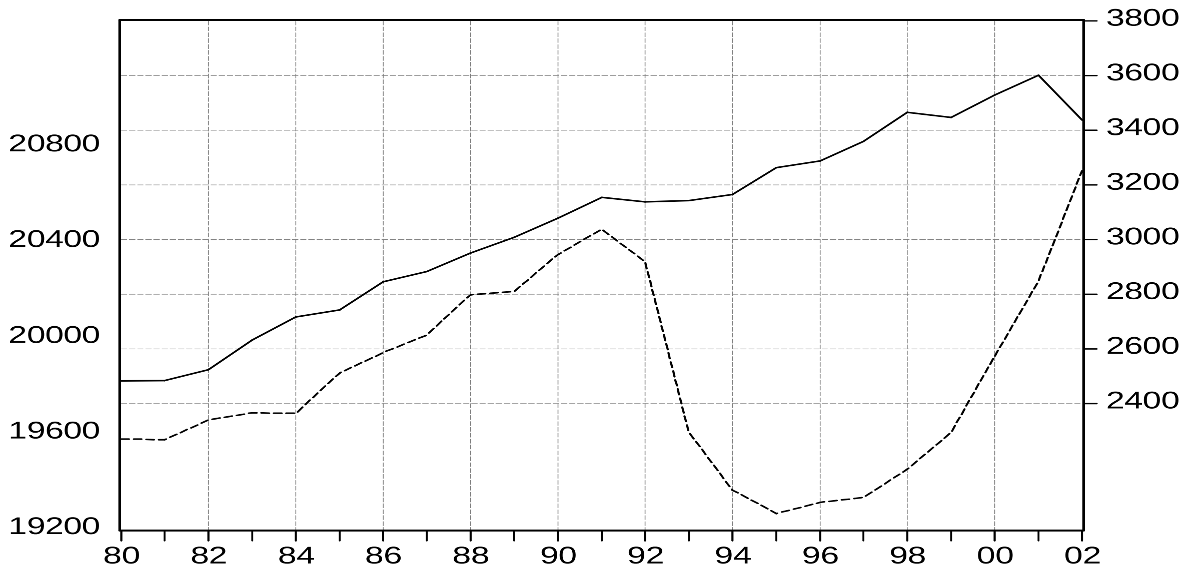
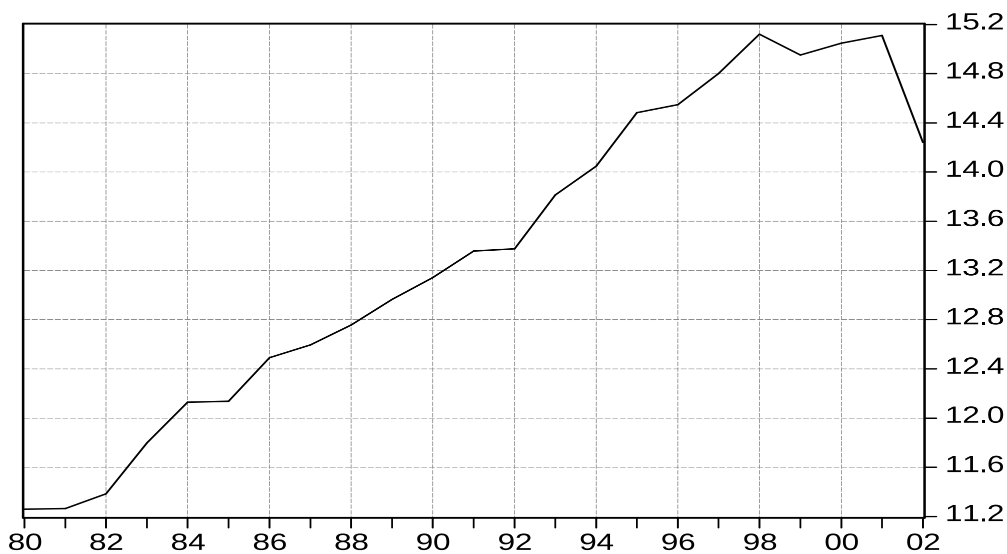


Figure 2. Irregularity Ratio - Dark Labour Input on Total (%) Italy 1980-2002 (Source: Istat)



APPENDIX 2. Bivariate VAR Residual Analyses. Sample 1980-2002.

Table 1a. VAR Model 1: D(FTE_reg); FTE_black. One lag. Eigenvalues (modulus): 0.91; 0.51

Multivariate tests			
Portmanteau Q-Stat 3 lags = 7.91 [0.4423] [a]	Normality J-B = 4.653 [0.3248]	Hetero X^2 No Cross Terms = 17.034 [0.3151]	Hetero X^2 Cross Terms = 21.25 [0.671]

D(x)=first difference of variable x; variables in logs; p-values in squared brackets; [a]=adjusted version. Point dummy in 1993 included.

Table 2a. VAR Model 2: FTE_reg; D(FTE_black). Two lags. Eigenvalues (modulus): 0.88; 0.45

Multivariate tests			
Portmanteau Q-Stat 3 lags = 8.763 [0.0673] [a]	Normality J-B = 4.553 [0.3364]	Hetero X^2 No Cross Terms = 28.93 [0.2228]	Hetero X^2 Cross Terms = 46.96 [0.2766]

See legend under table 1a.

Table 3a. VAR Model 3: D(FTE_reg); D(FTE_black). One lag. Eigenvalues (modulus): 0.58; 0.24

Multivariate tests			
Portmanteau Q-Stat 3 lags = 3.564 [0.8941] [a]	Normality J-B = 5.12 [0.2753]	Hetero X^2 No Cross Terms = 17.71 [0.1246]	Hetero X^2 Cross Terms = 23.75 [0.07]

See legend under table 1a. Point dummy in 2002 included.

Table 4a. VECM Model 4: VAR and cointegrating equation with intercepts (no trend): One lag.

Multivariate tests			
Portmanteau Q-Stat 3 lags = 9.81 [0.2787] [a]	Normality J-B = 3.921 [0.4168]	Hetero X^2 No Cross Terms = 17.39 [0.687]	Hetero X^2 Cross Terms = 26.15 [0.6676]

See legend under table 1a. Point dummies in 1993 and 2002 included.

Table 5a. VECM Model 5: VAR with intercept (no trend), cointegrating equation with intercept and trend. One lag.

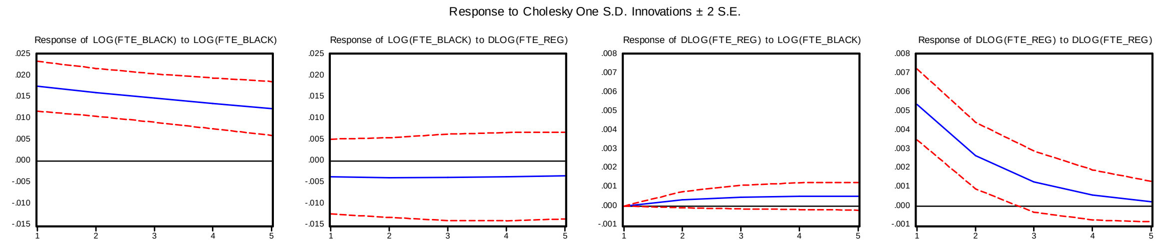
Multivariate tests			
Portmanteau Q-Stat 3 lags = [0.524] [a]	Normality J-B = [0.3248]	Hetero X^2 No Cross Terms = [0.3151]	Hetero X^2 Cross Terms = [0.671]

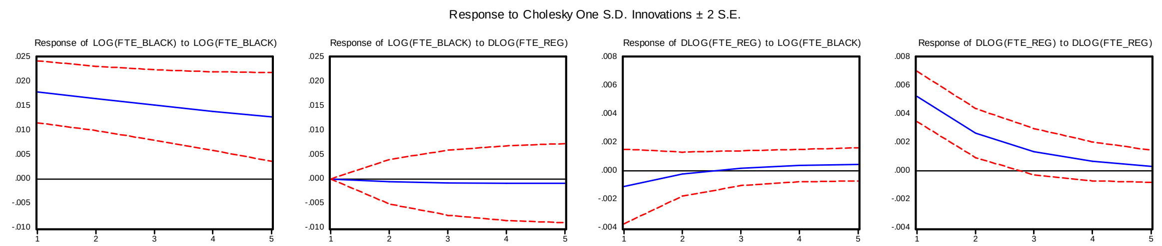
See legend under table 1a. Point dummies in 1993 and 2002 included.

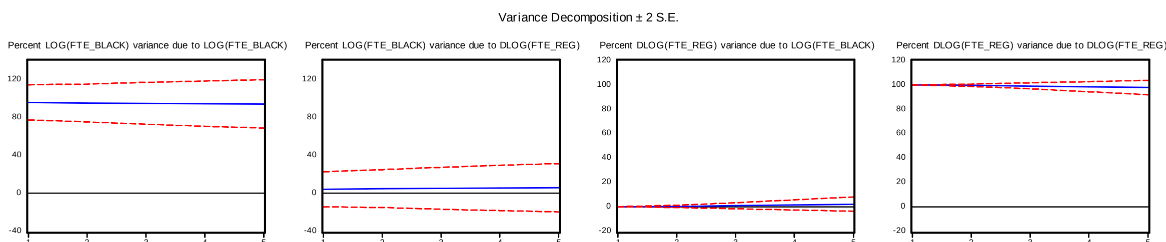
Appendix 3. Impulse Response and Innovation Accounting Analysis

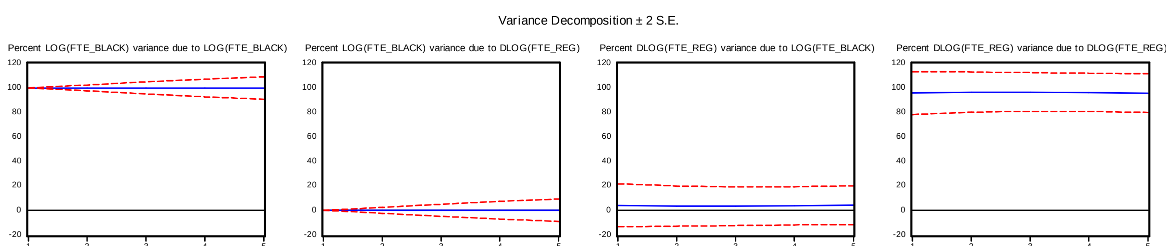
In all the figures, the Cholesky ordering for the relative impulse functions and for the variance decomposition analysis is FTE_reg; FTE_black => FTE_black; FTE_reg. In all figures, but the VECMs (for which S.E. can not be computed), the ± 2 S.E bands are drawn from 1000 Monte Carlo replications.

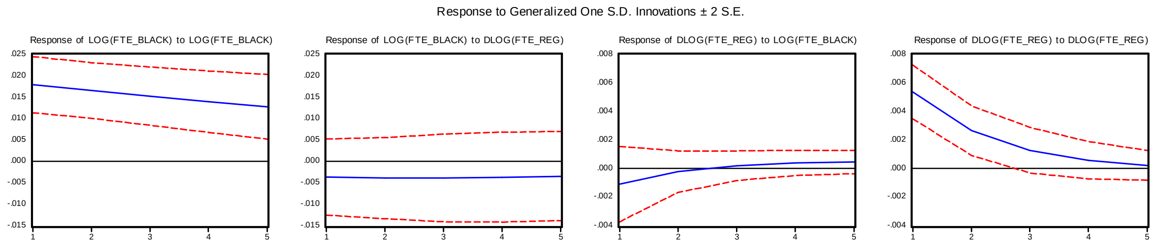
Model 1. Residual correlation: -0.20



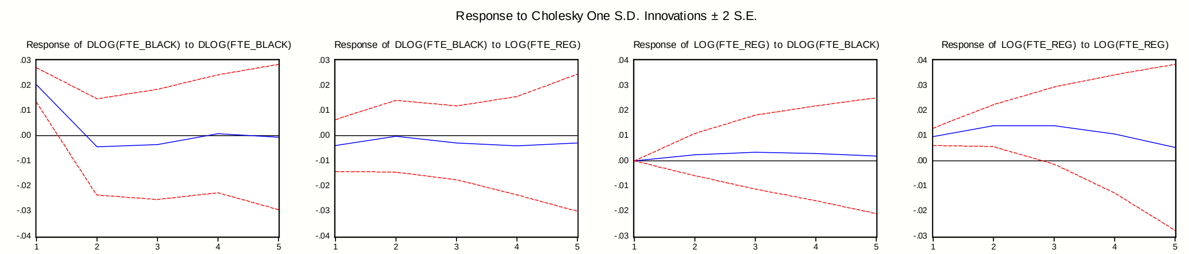


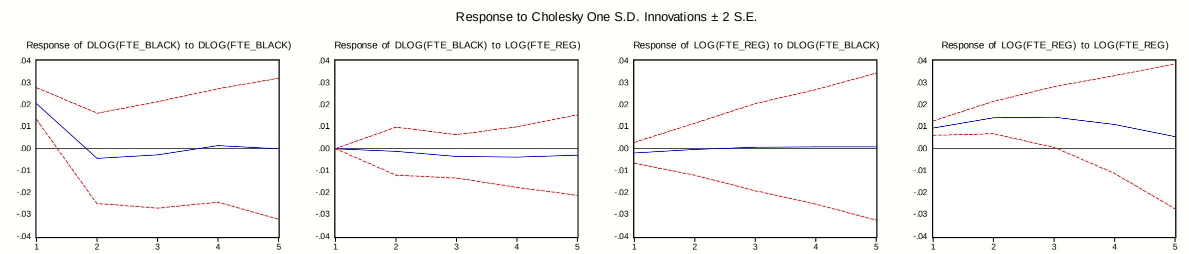


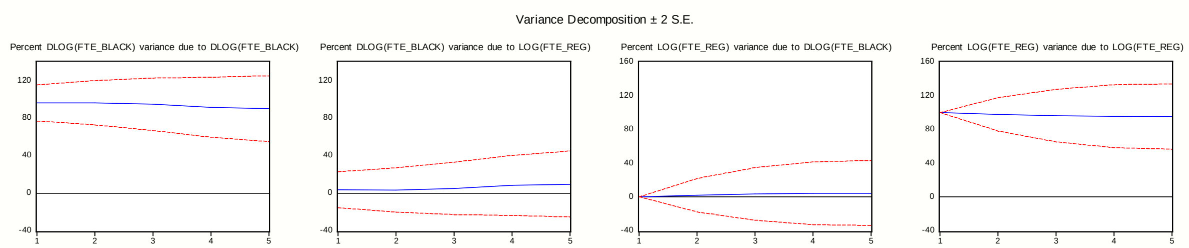


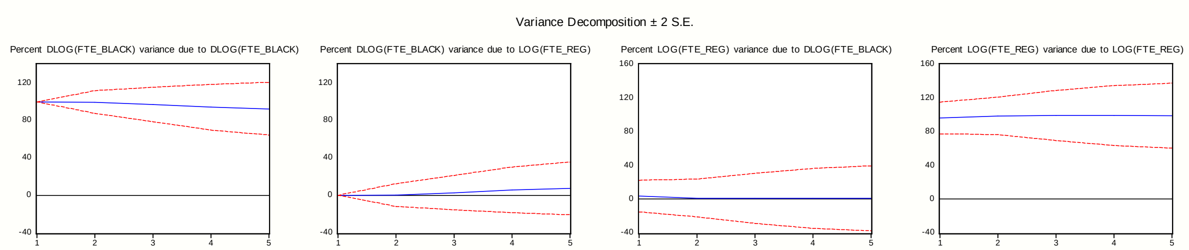


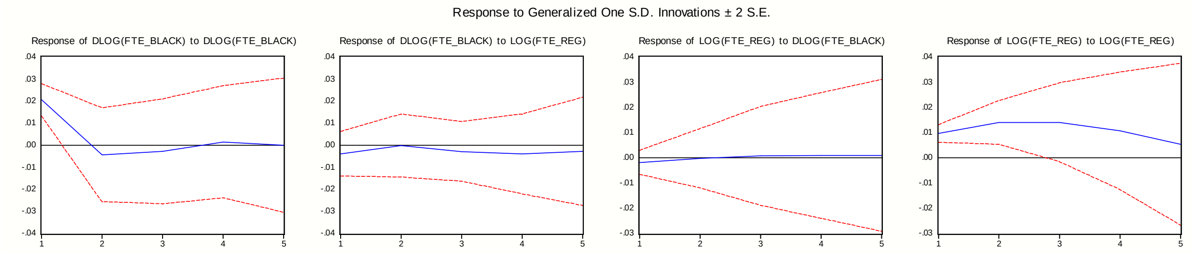
Model 2. Residual correlation: **-**0.18



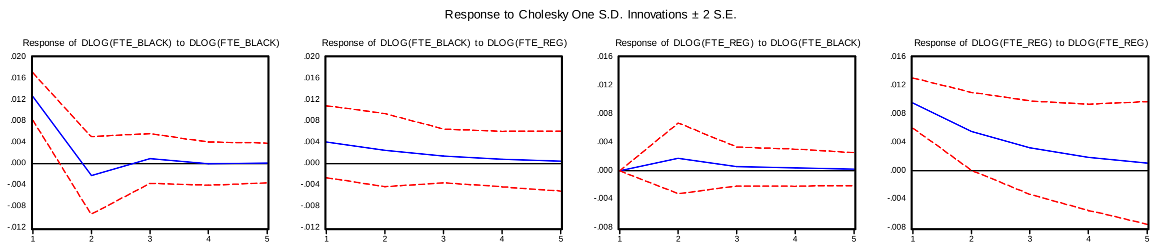


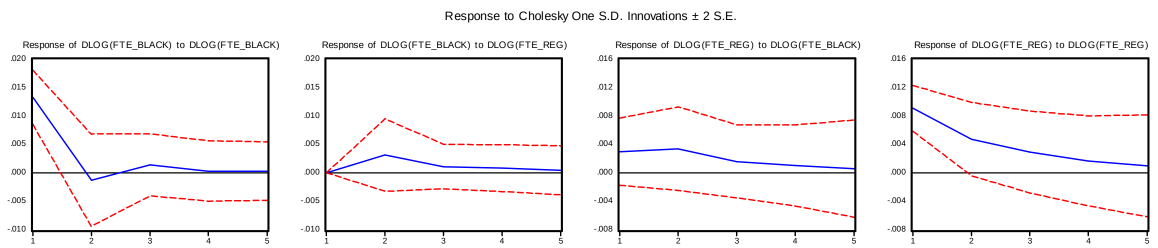


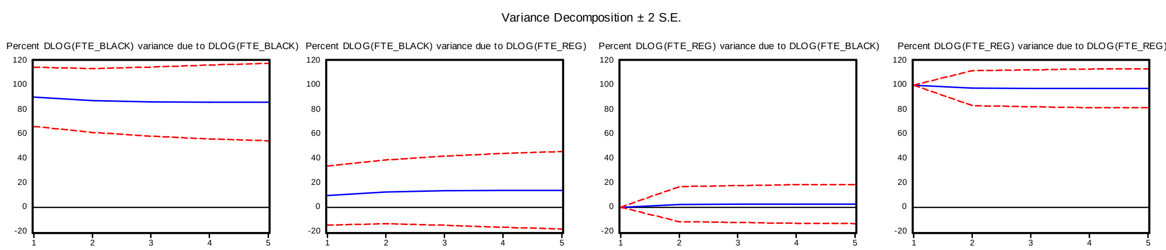


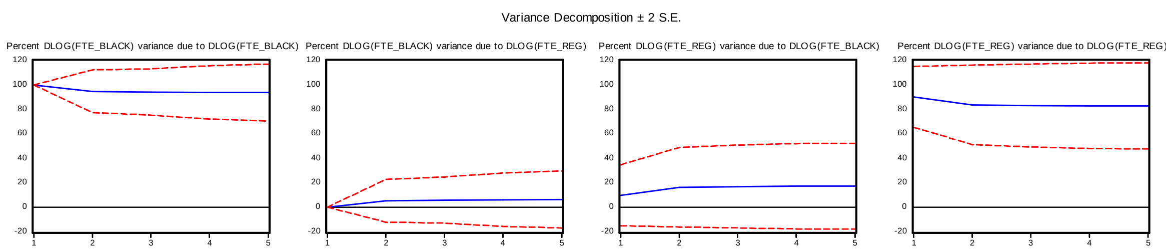


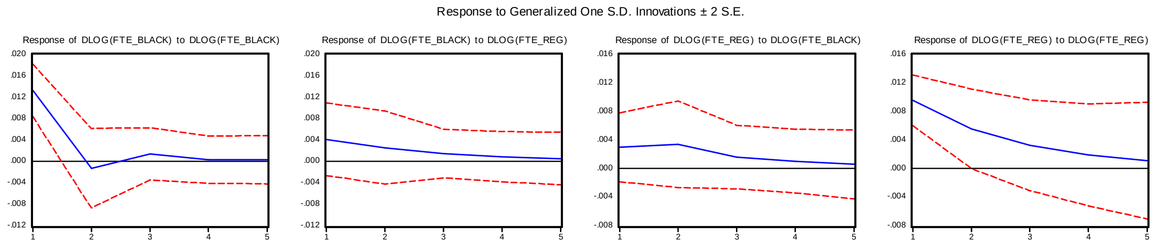
Model 3. Residual correlation: 0.31



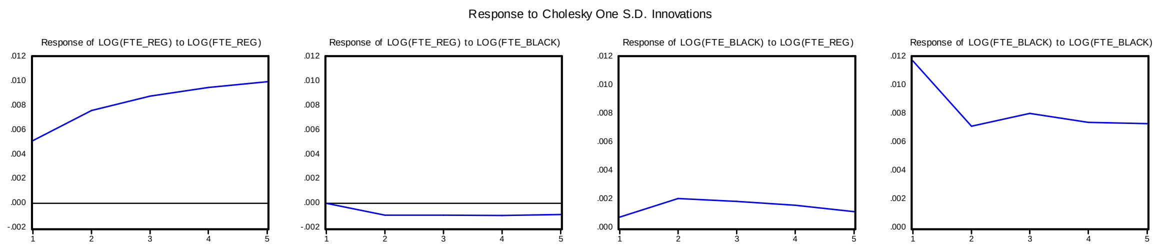


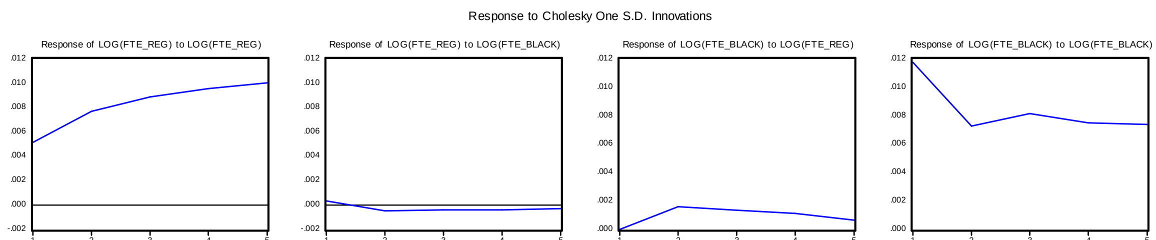




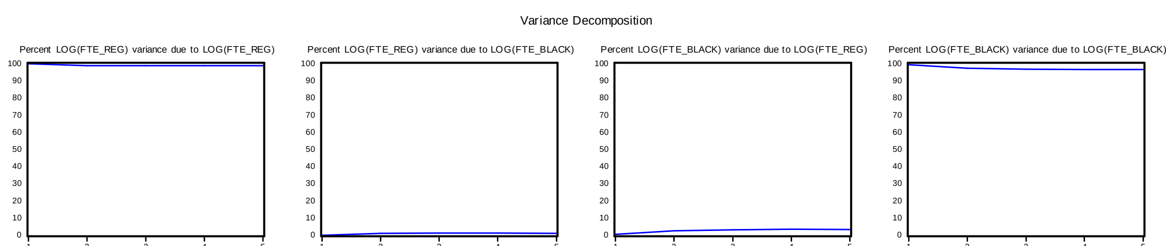


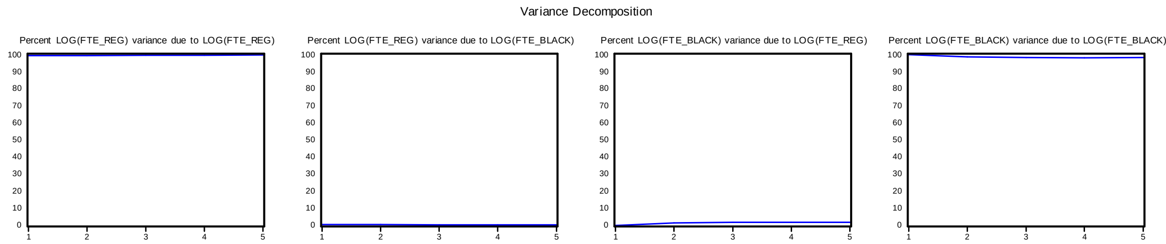
Model 4. Residual correlation: 0.06

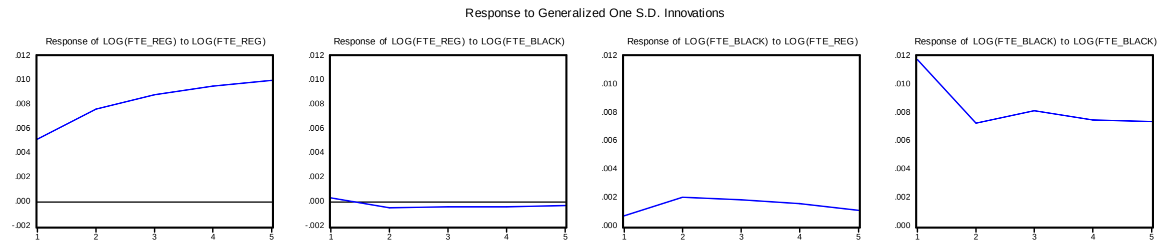




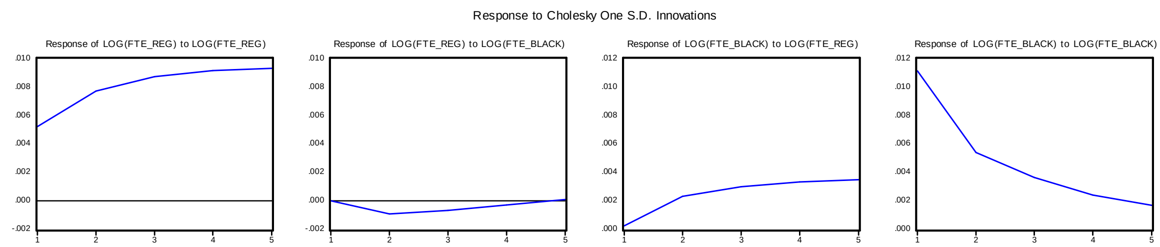
VEC Granger Causality/Block Exogeneity Wald Tests							
LHS: DLOG(FTE_REG)				**LHS: DLOG(FTE_BLACK)**			
Excluded	DLOG(FTE_BLACK)	χ^2	Prob.	Excluded	DLOG(FTE_REG)	χ^2	Prob.
		0.8235	0.3642			2.041	0.1531

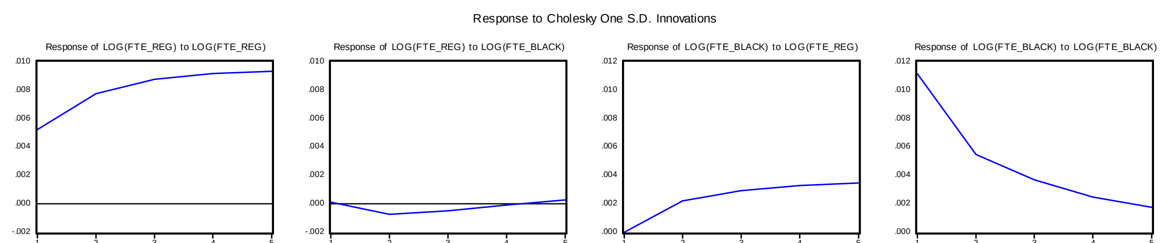






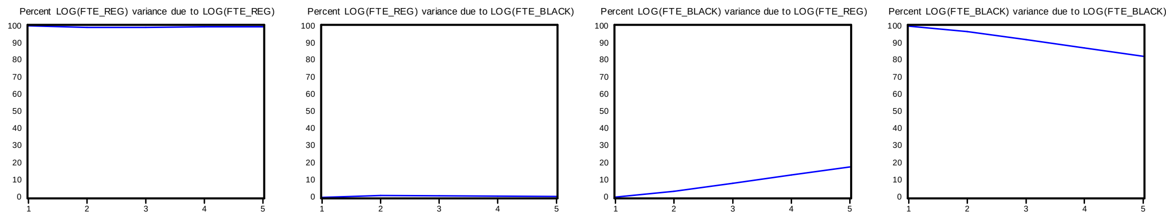
Model 5. Residual correlation: 0.02



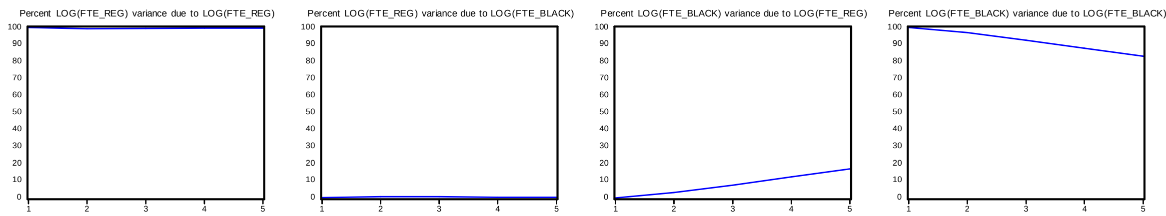


VEC Granger Causality/Block Exogeneity Wald Tests							
LHS: DLOG(FTE_REG)				**LHS: DLOG(FTE_BLACK)**			
Excluded	DLOG(FTE_BLACK)	χ^2	Prob.	Excluded	DLOG(FTE_REG)	χ^2	Prob.
		0.9432	0.3315			1.114	0.2912

Variance Decomposition



Variance Decomposition



Response to Generalized One S.D. Innovations