



An empirical examination of the determinants of the shadow economy

Journal:	<i>Applied Economics</i>
Manuscript ID:	AEL-2013-0558
Journal Selection:	Applied Economics Letters incorporating Applied Financial Economics Letters
Date Submitted by the Author:	30-Sep-2013
Complete List of Authors:	Acosta-González, Eduardo; Universidad de Las Palmas de Gran Canaria, Department of Quantitative Methods Fernández-Rodríguez, Fernando; Universidad de Las Palmas de Gran Canaria, Department of Quantitative Methods Sosvilla-Rivero, Simon; Universidad Complutense de Madrid, Department of Quantitative Economics
JEL Code:	C51 - Model Construction and Estimation < C5 - Econometric Modeling < C - Mathematical and Quantitative Methods, E26 - Informal Economy Underground Economy < E2 - Consumption, Saving, Production, Employment, and Investment < E - Macroeconomics and Monetary Economics, H26 - Tax Evasion < H2 - Taxation, Subsidies, and Revenue < H - Public Economics
Keywords:	Shadow economy, OECD countries, Tax evasion, GASIC

SCHOLARONE™
Manuscripts

An empirical examination of the determinants of the shadow economy

Eduardo Acosta-González^a

Fernando Fernández-Rodríguez^a

Simón Sosvilla-Rivero^{b*}

Abstract

Using a statistical methodology guided only by data and based on a genetic algorithm, we select the best econometric model for explaining the determinants of the size of the shadow economy, its main determinants being: taxes on capital gains of individuals, corporate taxes on income, profits and capital gains, domestic credit, bank secrecy, ethnic fractionalization, urban population, globalization, corruption and the socialist legal origin of country.

JEL Codes: C51, E26, H26

^a Department of Quantitative Methods, Universidad de Las Palmas de Gran Canaria, Campus de Tafira, 35017 Las Palmas de Gran Canaria, Spain

^b Department of Quantitative Economics, Universidad Complutense de Madrid, Campus de Somosaguas, 28223 Madrid, Spain

* Corresponding author. E-mail: sosvilla@ccee.ucm.es

I. Introduction

The determinants of the shadow economy remain controversial (see Schneider and Enste, 2000, for a survey).

Using the massive database gathered by Thiessen (2010), the main contribution of our paper consists of selecting the determinants of the size of the shadow economy from a set of 274 variables suggested in the theoretical and empirical literature as potential causes and influences. We employ a purely statistical methodology which automatically selects the factors in the econometric model in a process guided only by data. Our methodology properly handles multicollinearity problems in the estimations that could arise due to the redundancy of the information provided by the factors. Besides, it offers a parsimonious model of the shadow economy, containing few factors with as much information as possible that captures the essential characteristics of the data, and can help in identifying the main reasons explaining its size.

II. Methodology

In order to avoid data mining problems for constructing the best econometric model explaining the empirical determinants of the size of the shadow economy, we follow the general to specific approach based on the theory of reduction (Hendry, 1995). It means that given an endogenous variable Y , which represents the indicator of the shadow economy, and a set X_1, \dots, X_K of 274 potential factors explaining it, the problem that we face is finding the best sub-model of the form:

$$Y = \beta_0 + \beta_1 X_{i_1} + \dots + \beta_k X_{i_k} + \varepsilon,$$

where $\{i_1, i_2, \dots, i_k\} \subseteq \{1, 2, \dots, K\}$, $k \leq K$, being k the number of factors in the model, excluding the intercept.

1
2
3
4
5 The problem is that the number of sub-models where the criteria developed by
6 econometric theory for selecting the best model is 2^K . So, for $K = 274$, the number of
7 possible models is higher than 10^{135} . In order to resolve this intractable problem, we
8 follow the heuristic strategy proposed by Hoover and Perez (1999) by searching all the
9 feasible reduction paths when simplifying the general model. In particular, we use the
10 methodology developed in Acosta-González and Fernández-Rodríguez (2007). This is a
11 procedure for automatic selection of factors in the model, guided only by the data,
12 which is carried out using a Genetic Algorithm where the lost function is the Schwarz
13 Information Criterion (SIC, henceforth). This methodology avoids the tendency to over-
14 identify models detected in several popular heuristic methods for selecting models [like
15 stepwise, Lovell (1983)] and, by allowing all possible combinations of potential
16 determinants, renders results that are robust to any specification.
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33

34 Therefore, starting with a large number of potential causes and an indicator of the
35 shadow economy considered in the literature to take into account all the information
36 about its possible determinants, our methodology selects the best econometric model, in
37 the sense of the SIC.
38
39
40
41
42
43
44

45 The SIC advises choosing the econometric model which minimizes the expression
46

$$SIC = \log \hat{\sigma}^2 + c \frac{k+1}{N} \log(N) \quad (1)$$

47
48
49
50
51
52
53 where $\hat{\sigma}^2 = \frac{e'e}{N}$ is the variance of the residuals e , N is the sample size, and k is the
54 number of factors in the model, excluding the intercept. The correcting factor c avoids
55
56
57
58
59
60

1
2
3 the possibility of over-parameterized models and solves the trade-off between the in
4 sample goodness of fit and the out of sample forecasting ability. The higher the value of
5 c , the higher the penalty for the introduction of more factors in the model.
6
7
8
9

10 11 12 **III. Data and empirical results**

13
14 We use the database gathered by Thiessen (2010) for 38 countries for the period 1991-
15 2007¹. Given the high amount of missing data present in the Thiessen's database, in
16 order to produce a data matrix without missing values, we have employed a careful case
17 wise deletion procedure, guided by our genetic algorithm, to eliminate the missing
18 observations and variables. From the initial set of 646 observations [17 years times 38
19 countries], the model was finally estimated with 371 observations.
20
21
22
23
24
25
26
27

28
29 Regarding the indicator of the size of the shadow economy, we use money cash holding
30 relative to M2. As for the explanatory variables, the database includes 294 potential
31 determinants of the shadow economy covering such categories as constitutional system,
32 tax and social security burden, administrative burden, quality of administration and
33 justice system, economic institutions, values and moral aspects, and other influences
34 and subjective factors (such as globalization, income inequality, ageing, etc.)².
35
36
37
38
39
40
41
42
43
44

45 Table 1 reports the results provided by our genetic algorithm in order to search for the
46 best set of variables explaining the shadow economy. As can be seen, we obtain a
47
48
49

50
51 ¹ We have data on 30 OECD countries (Australia, Austria, Belgium, Canada, Cyprus,
52 Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, Latvia,
53 Lithuania, Luxembourg, Malta, Mexico, Netherlands, New Zealand, Norway, Portugal,
54 South Korea, Spain, Sweden, Switzerland, Turkey, United Kingdom and United States)
55 and 8 Eastern European countries (Bulgaria, Czech Republic, Estonia, Hungary, Poland,
56 Romania, Slovak Republic, and Slovenia).

57 ² See Thiessen (2010, Appendix 1) for a detailed account of the variables included in
58 each of these seven major groups of potential determinants.
59
60

1
2
3 negative and significant association between the shadow economy and taxes on capital
4 gains of individuals, indicating that informality is a form of tax evasion. This is in line
5 with previous theoretical and empirical analyses (e.g. Constantinides 1982, Stiglitz
6 1983, Poterba 1987, OECD, 2006) that have suggested that there exist considerable
7 opportunities for avoiding tax and for generating capital losses to offset ordinary
8 income (especially for higher incomes), being therefore a substitute for the shadow
9 economy as a means of evading taxes.
10
11

12
13
14 We find a significant positive association between the shadow economy and taxes on
15 income, profits and capital gains (as a percentage of GDP), suggesting that these have a
16 negative impact on the decisions made by economic agents on whether to operate
17 officially, driving entrepreneurs and businesses underground³.
18
19

20
21
22 The results of the regressions in Table 1 show that domestic credit (as a percentage of
23 GDP) significantly influences the size of the hidden economy. This is consistent with
24 the fact that although domestic credit is positively and significantly correlated with
25 business and entry density rates, due to burdensome regulations, high marginal tax rates,
26 the absence of monitoring and compliance (of both registration and tax regulations), and
27 other weaknesses in the business environment, many firms might find it optimal to
28 evade regulations and operate in the informal sector (Klapper et al., 2008).
29
30
31

32
33
34 A strongly significant negative influence is also shown with regard to the bank secrecy
35 dummy (1 = bank secrecy law is relatively effective; 0 = no bank secrecy law or not
36 effective). This may indicate that bank secrecy would be a substitute for shadow
37
38
39
40
41
42
43
44
45
46
47
48

49
50
51
52
53
54
55
56
57 ³ See Friedman *et al.* (2000) for a simple model of an entrepreneur's decision to operate
58 officially or unofficially depending on corporate taxes
59
60

1
2
3 economy by facilitating illicit financial flows worldwide for money laundering. Indeed,
4
5 the world's biggest players in the supply of financial secrecy are rich nations (see,
6
7 Walker and Unger, 2009).
8
9

10
11 The ethnic fractionalization index shows a significant negative influence on the size of
12
13 the informal economy, in line with Montalvo and Reynal-Querol (2005), who document
14
15 that a high degree of ethnic fractionalization has negative effects on investment and
16
17 increases rent seeking activities.
18
19

20
21 With regard to urban population, our result is consistent with that of Elgin and Oyvaty
22
23 (2013), who argue that although the share of the informal sector grows in the early
24
25 phases of urbanization due to several pull and push factors (see. Smith, and Scarpaci,
26
27 2010), it tends to decrease in the latter phases of urbanization.
28
29
30

31
32 Concerning globalization, increasing competition in goods and factor markets raises
33
34 uncertainty of income and employment, increasing the size of the shadow economy
35
36 (Karlinger, 2009). Moreover the development of the networking-based economy has
37
38 given birth to a large group of shadow enterprises, making the concept of national
39
40 boundaries and distance for certain economic activities meaningless (Shangquan, 2000).
41
42
43
44
45

46
47 In relation to corruption, our finding supports the view that corruption and the shadow
48
49 economy are complements—in high income countries and thus, a substitute for the
50
51 official economy as suggested by Johnson *et al.* (1997) and Hindriks *et al.* (1999)⁴.
52
53
54

55
56
57 ⁴ To facilitate the interpretation, we have multiplied the original variable in Thiessen's
58
59 database by -1, so a higher value of this variable indicates greater corruption.
60

1
2
3 Finally, we detect a significant and positive influence for a dummy indicating that the
4 legal origin of the country is socialist, suggesting that the legacy of high degree of
5 bureaucratic discretion; undeveloped market institutions and enforcement mechanisms,
6 and low degree of economic liberalization would have adversely influenced formal
7 activity (Eilat and Zinnes, 2002).
8
9
10
11
12

13
14
15
16 Turning to the goodness of fit, the value of the R-squared statistic reported in Table 1
17 suggests the relative success of our regression model in predicting the shadow economy
18 within the sample.
19
20
21
22
23

24
25 [Table 1]
26
27
28

29 **IV. Conclusion**

30
31 We have empirically examined the determinants of the size of the shadow economy
32 using a statistical methodology guided only by data and based on a genetic algorithm.
33
34 We find a positive influence for corporate taxes, domestic credit, socialist legal origin of
35 the country and corruption, as well as a negative influence for taxes on capital gains,
36 bank secrecy, ethnic fractionalization, urbanization and globalization.
37
38
39
40
41
42
43
44

45 **Acknowledgements**

46
47 The authors are very grateful to Ulrich Thiessen for kindly providing the dataset.
48
49 Financial support from the Spanish Institute for Fiscal Studies (project IEF 13/2013) is
50 also gratefully acknowledged.
51
52
53
54
55
56
57
58
59
60

References

Acosta-González, E. and Fernández-Rodríguez, F. (2007) Model selection via genetic algorithms illustrated with cross-country growth data, *Empirical Economics*, **33**, 313-337.

Constantinides, G. M. (1984) Optimal stock trading with personal taxes: implications for prices and the abnormal January returns, *Journal of Financial Economics*, **13**, 65-89

Elgin, C. and Oyvaty, C. (2013) Lurking in the cities: urbanization and the informal economy, *Structural Change and Economic Dynamics*, **27**, 36-47.

Eilat, Y., Zinnes, C. (2002) The shadow economy in transition countries: friend or foe? A policy perspective, *World Development*, **30**, 1233–1254.

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*, **76**, 459–493

Hendry, D.F. (1995) *Dynamic Econometrics*, Oxford University Press, Oxford.

Hindriks, J.; Muthoo, A. and Keen, M. (1999) Corruption, extortion and evasion, *Journal of Public Economics*, **74**, 395-430.

Hoover, K.D. and Perez, S. J. (1999) Data mining reconsidered: encompassing and the general- to-specific approach to specification search, *Econometrics Journal*, **2**, 167-191.

1
2
3 Johnson, S., Kaufmann, D. and Shleifer, A.(1997) The unofficial economy in transition,
4
5 *Brookings Papers on Economic Activity*, **28**, 159-221.
6
7

8
9
10 Karlinger, L. (2009) The underground economy in the late 1990s: evading taxes, or
11
12 evading competition? *World Development*, **37**, 1600-1611.
13
14

15
16 Klapper, L., Amit, R. and Guillén, M. F. (2008) Entrepreneurship and firm formation
17
18 across countries, in *International Differences in Entrepreneurship*, L. Lerner and A.
19
20 Schoar (Eds.), National Bureau of Economic Research, Cambridge, MA
21
22

23
24
25 Lovell, M.C. (1983). Data mining, *Review of Economics and Statistics*, **65**, 1-12.
26
27

28
29
30 Montalvo, J. G. and Reynal-Querol, M. (2005) Ethnic diversity and economic
31
32 development, *Journal of Development Economics*, **76**, 293–323
33
34

35
36 OECD (2006) *Taxation of capital gains of individuals: policy considerations and*
37
38 *approaches*. OECD, Paris.
39
40

41
42
43 Poterba, J. M. (1987) How burdensome are capital gains taxes?: evidence from the
44
45 United States, *Journal of Public Economics*, **33**, 157–172.
46
47

48
49
50 Schneider, F., and Enste, D. H. (2000) Shadow economies: size, causes, and
51
52 consequences, *Journal of Economic Literature*, **38**, 77–114
53
54
55
56
57
58
59
60

1
2
3 Shangquan, G. (2000) Economic globalization: trends, risks and risk prevention, CDP
4
5 Background Paper 001, Department of Economics and Social Affairs, United Nations.
6
7

8
9
10 Smith, D. W. and Scarpaci, J. L. (2010) Urbanization in transitional societies: an
11
12 overview of Vietnam and Hanoi, *Urban Geography*, **21**, 745-757
13

14
15
16 Stiglitz, J. E. (1983) Some aspects of the taxation of capital gains, *Journal of Public*
17
18 *Economics*, **21**, 257-294.
19

20
21
22
23 Thiessen, U. (2010) The shadow economy in international comparison: options for
24
25 economic policy derived from an OECD panel analysis, *International Economic*
26
27 *Journal*, **24**, 481-509.
28

29
30
31 Walker, J. and Unger, B. (2009) Measuring global money laundering, *Review of Law*
32
33 *and Economics*, **2**, 821–853.
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

TABLE 1. Estimation results for a correcting factor $c=2$ in the Schwarz Information Criterion

	Coefficient	Std. Error	t-Statistic	Prob.
Constant	184.2410	22.7739	8.0900	0.0000
Taxes on capital gains of individuals (% of GDP)	-129.9080	10.6907	-12.1515	0.0000
Corporate taxes on income, profit & c.g. (% of GDP)	8.2846	1.4892	5.5631	0.0000
Domestic credit (% of GDP)	1.1084	0.0352	31.5224	0.0000
Bank secrecy dummy ¹	-39.3740	5.4486	-7.2264	0.0000
Ethnic fractionalization Index	-174.5049	21.6672	-8.0539	0.0000
Urban population (% of total)	-1.9017	0.2645	-7.1890	0.0000
Globalization index	-1.1254	0.2939	-3.8294	0.0002
Corruption	8.3117	2.0670	4.0206	0.0001
Legal origin of country, Socialistic	38.4636	8.4153	4.5707	0.0000
R-squared	0.8119	F-statistic		173.1621
Adjusted R-squared	0.8073	Prob(F-statistic)		0.0000
Standard Error of regression	39.5097	Schwarz criterion		24.1386