Savings and the informal sector

by

Stephen Dobson^

Carlyn Ramlogan-Dobson^

Eric Strobl*

^Hull University Business School, University of Hull, Hull, UK

*Department D'Economie, Ecole Polytechnique, Palaiseau Cedex, France

Abstract

In many countries the informal sector is a vital source of employment and income. But little is known about the impact of this sector on savings, which are crucial in promoting investment and growth. This paper finds an inverse relationship between savings rates and the informal sector when the informal sector is small. Once the informal sector reaches a certain size, further growth in the size of the informal sector boosts savings rates. The non-linear relationship is confirmed in both parametric and semi-parametric estimations. Rather than allowing the informal sector to grow unchecked, policy should focus on removing barriers for successful operation of business in the formal sector.

Key words: informal sector, savings rates, semi parametric

This is an Accepted Manuscript of an article published by Taylor & Francis in International Review of Applied Economics on 7th January 2020, available online: https://www.tandfonline.com/10.1080/02692171.2019.1707783

1. Introduction

The empirical literature on savings is underpinned by two key theories: the life cycle hypothesis (Modigliani and Brumberg, 1954) and the permanent income hypothesis (Friedman, 1953). Since both theories have limitations within empirical applications, numerous factors are included as determinants of savings rates in savings studies (e.g., Deaton, 1991; Carroll, 1997; Tsikata, 1998; Loayza et al. 2000; Smith, 2001; Disney, 2006; Horioka and Hagiwara, 2011; Beckmann et al. 2013). One factor that is missing in empirical studies of savings is the informal sector. This paper fills an important gap in the literature by examining the relationship between the informal sector and savings rates in a large sample of countries.

One possible reason for the absence of the informal sector in savings studies is limited data availability until more recent times. A second, and more likely reason, is the 'conventional wisdom' that says the informal sector is a transitional feature of economies, especially developing ones. This view has been challenged in places like Latin America and Africa, where the informal sector seems to be embedded in the culture as people readily accept it as an integral part of society (Bekkers and Stoffers, 1995; Charmes, 2000). Informal sector employment is a positive choice for a significant proportion of the labor force and not necessarily a temporary situation while waiting for a job in the formal sector (Feiss et al. 2010).

Research also challenges the notion that the informal sector is full of low paid workers. Badaoui et al. (2008) find no pay discrepancy between the formal and informal sector for top-tier individuals in South

Africa while Bargain and Kwenda (2010) note that earnings differentials across the formal and informal sectors are noticeably small in Brazil. Gelb et al. (2009) find the productivity of informal firms is almost indistinguishable to that of formal firms in East Africa while Mukerjee (2004) shows that, with few exceptions, productivity in the informal sector in India has been increasing. Informal sector activity has also been approved by government. In Peru, for example, the government recognized informal sector minibuses as the main means of transportation of the masses (De Soto, 1989). Indeed, in many countries in Latin America and the Caribbean the informal sector is the mainstay of public transportation because official public transport is inadequate and unreliable in parts of the world.

Research has shown the informal sector to be relatively large in many countries. Schneider et al. (2010) estimate the average size of the informal sector (as a percentage of GDP) over the period 1999-2007 for 161 countries. By sorting countries into regions, as defined by the World Bank, Latin America and the Caribbean is shown to have the highest value of the informal sector (41 per cent), followed by Sub-Saharan Africa (40 per cent) and then Europe and Central Asia (39 per cent). The lowest estimate is for high income OECD countries (13 per cent). Data for individual countries shows that the informal sector accounts for over 40 per cent of GDP in 50 countries, with four of those countries having values of over 60 per cent (see also Schneider, 2012). Table 1 in Section 3 reports summary statistics on informal sector size. Schneider (2012) also measures informality in terms of the share of informal employment in total (non-agricultural) employment for the period 2000-2007. The highest average values are in Asia (70 per cent) and Africa (61 per cent). Laoyza (2016) notes that the typical developing country employs over 70 per cent of its labour force informally and produces about 30 per cent of its output informally.[1]

In view of the questions raised about the conventional notion of the informal sector and given the relatively large size of the informal sector in many countries, it seems reasonable to hypothesize that the informal sector may be an important factor in determining savings rates. This idea is reinforced by examining a plot of savings rates against the size of the informal sector, as shown in Figures 1 and 2. A non-linear relationship is evident in Figure 1, while a clear positive link is shown in Figure 2 once the sample is restricted to those countries with a relatively large informal sector (more than 38 per cent of GDP).

We estimate the relationship between savings rates and the informal sector using parametric and semi-parametric econometric methods. Our results show a U-shaped relationship: when the informal sector is small savings rates fall with growth in the size of the informal sector but once the informal sector reaches a certain size further growth in the informal sector leads to an increase in savings rates. In general, the turning point occurs when the informal sector accounts for around 38-40 per cent of GDP. This indicates that countries with a large informal sector have an opportunity to mobilize savings to a much greater extent than conventional wisdom would suggest. These results also suggest that previous work on the determinants of savings that include countries with a large informal sector may be misspecified.

The structure of the paper is as follows. Section 2 considers the conceptual link between savings and the informal sector. The empirical model and data are discussed in Section 3. Results are presented and analysed in Section 4. In Section 5 we examine a further test of robustness of the results and Section 6 concludes.

2. Informality, savings and output

There are various definitions of the informal economy in the literature. According to Feige (1994) the informal sector comprises all currently unregistered economic activities that contribute to the official calculation of GNP. Smith (1994, p. 18) defines the informal sector as 'market-based production of goods and services, whether legal or illegal, that escapes detection in the official estimates of GDP'. This broad notion of the informal sector has been expressed differently by a number of authors as economic activities (and the income derived from them) that avoid government regulation, observation and taxation (see, for example, Feige, 1989; Dell'Anno and Schneider, 2003). A narrower definition is used by Schneider (2012) whereby the informal sector includes all legal market-based production that contributes to national income but escapes detection in official estimates. This definition therefore excludes illegal underground economic activities (such as burglary, robbery and drug dealing) as well as informal household production. We prefer the narrow definition of the informal sector because Schneider's estimates of the size of informal sector, which are used in our empirical estimations, are based on the narrow definition. Also, data based on the narrow definition provides the largest coverage of countries.

One way of estimating the size of the informal sector is by direct methods such as household surveys or tax audits. These tend to underestimate informality as they do not include individuals engaged in erratic bouts of informal employment or very small-scale activities. An alternative approach is to use indirect methods. These use discrepancies between official estimates of macroeconomic variables to approximate a gap in economic activity (e.g. electricity consumption). This approach is not without problems as not all informal activity requires the exchange of cash or the use of energy. A third

approach to estimating the size of the informal sector is to use statistical methods. These are used to construct multiple-indicators multiple causes models (MIMIC), which treat the informal sector as a latent variable that can be estimated using quantitative indicators reflecting unregistered activities. One limitation of this approach is that the predicted values are sensitive to the underlying model specification.

The informal sector may impact savings rates directly and indirectly. Informal sector income is often transient and volatile (Browning and Lusardi, 1996; Granada and Hamann, 2015), which encourages precautionary saving. The direct impact fits with the life cycle hypothesis and the permanent income hypothesis since both imply that transitory income will be saved.[2] Aside from uncertain or transient income, an absence of unemployment benefit, health insurance and pension schemes within the informal sector may encourage greater savings.[3] Also, informal sector entrepreneurs often face restrictions in accessing credit, which may boost savings as a means to facilitate start-ups or further business investment.

The indirect impact of the informal sector on savings operates through output (income). In countries with poor institutions and poor governance, the informal sector allows entrepreneurs to pursue ventures which would not be possible in the formal economy due to high levels of bureaucracy, regulation and rent seeking. In effect, firms become victims of predatory behavior by corrupt officials. To the extent that the informal sector circumvents 'bad laws' and reduces the costs of regulation and bureaucracy, informal production can substitute for that in the formal sector, leading to an increase in output (e.g., Sarte, 2000; Djkanov et al. 2002; Choi and Thum, 2005; Dreher et al. 2009; Elgin and Uras, 2013). In countries with robust institutions and good governance activity in the informal sector is considered to

be less productive than in the formal sector for reasons relating to size, access to credit, training, innovation and property rights. Furthermore, because informal firms escape taxation they undermine the ability of government to provide public goods and effective institutions, which further harms output (e.g., Johnson et al. 1997; Kauffman et al. 1997; Docquier et al. 2016).[4] A formal presentation of the link between the informal sector and output is in the Appendix.

In countries with poor institutions (large informal sector) the direct and indirect impact work together to boost savings rates: savings increase for precautionary reasons and because of higher output. In countries with strong institutions (small informal sector) the direct impact is small and is outweighed by the indirect impact of lower output, leading to a fall in savings rates. This is suggestive of a possible non-linear relationship between the size of the informal sector and savings rates. This idea is supported by informal evidence of a plot of savings rates against the size of the informal sector, as shown in Figures 1 and 2.

In summary, in countries with poor institutions and heavy regulation bureaucrats maximise rent seeking gains. Individuals wishing to start a business are often subject to insurmountable delays unless bribes are paid. If entrepreneurs operate informally the influence of corrupt agents is reduced and informal production substitutes for that in the formal sector. Production is further facilitated as firms develop a network of contacts over time (Cule and Fulton, 2009). As networks become stronger it becomes less expensive to operate informally, thus leading to further increases in output and savings.

3. Data and empirical issues

In the first instance we estimate a standard savings equation for a broad cross-section of developed and developing countries:

$$S_i = \beta_0 + \beta_1 I_i + \beta_2 I_i^2 + \alpha_i Z_i + u_i \quad (1)$$

Where S_i is the savings rate for country i, I is the size of the informal sector, Z is a vector of other explanatory variables and u is the error term. Together I and I^2 capture the non-linear impact of the informal sector on savings rates. Each variable is averaged over the period 1999-2007 and the estimation method is OLS.

This paper measures the informal sector in terms of percentage of 'official' GDP using data from Schneider et al. (2010) (and reproduced in Schneider, 2012). We employ two measures of savings: (i) gross national savings = gross national disposable income − final consumption expenditure and (ii) gross domestic savings = gross domestic disposable income − final consumption expenditure. In both (i) and (ii) savings is equal to public plus private savings.

The variables in Z are commonly used in the savings literature: dependency ratios, inflation, real effective exchange rate, rate of interest, money to GDP ratio, remittances, domestic credit to the private sector as a ratio of GDP, government budget surplus/deficit, inequality, corruption, openness, political stability, and income. Inequality is measured with the Gini coefficient and data are drawn from the United Nations World Income Inequality Database (WIID) (UNU-WIDER, 2005).[5] We use several definitions of corruption: international country risk guide (ICRG) corruption index, the corruption perception indeed (CPI), and the IMF's governance indicator. Data on GDP and openness are obtained from Penn World Tables (PWT7).[6] The polity data base is used for data on political stability. Data on

the remaining explanatory variables are obtained from World Development Indicators (2010). Table A1 in the Appendix reports variable definitions. Summary statistics are shown in Table 1.

One of the main problems in estimating equation (1) is the potential endogeneity of the explanatory variables. Instrumental variable estimation is often used to overcome this difficulty but there are problems in obtaining reliable instruments for all potentially endogenous controls. Therefore, we experiment with lagged values of the explanatory variables. For this procedure data are divided into two sub periods, 1999-2003 and 2004-2007, with the dependent variable measured as its average over the second sub period and the explanatory variables as averages over the first sub period (lagged variables). We also experiment with several other cross section models and a panel data model in order to confirm the robustness of the results.

4. Results and discussion

The empirical estimates of equation (1) are presented in Tables 2 and 3. The baseline model is shown in column (1) of Table 2. The informal sector variable is not significant.[7] However, when we include a quadratic term, so as to capture the non-linear impact, both the level and the squared terms are statistically significant. Moreover, the signs indicate that the relationship between the informal sector and savings can be explained by a U-shape: when the informal sector is small savings and the informal sector are inversely related but beyond a certain size further growth in the informal sector is associated with higher savings. This result is consistent with the discussion in Section 2. The coefficients suggest that the turning point in the relationship occurs when the size of the informal sector is equivalent to 38 per cent of GDP. This represents 64 countries in our sample, all of them developing countries. In other words, countries with a relatively large informal sector see an increase in savings rates when the

informal sector grows larger, *ceteris paribus*. This result is consistent with other work on informality and economic outcomes (not savings) (e.g., Gelb et al. 2009; McKenzie and Sakho, 2010; Benjamin and Mbaye, 2012; Dobson and Ramlogan-Dobson, 2012).

To investigate the robustness of this result we examine other models in Tables 2 and 3. In columns (2) and (3) of Table 2 we split the data into two time periods: 1999-2003 (sub sample 1) and 2004-2007 (sub sample 2). A non-linear relationship between the informal sector and savings rates is obtained in these models. Also, the turning point is similar to that found in column (1). Column (4) of Table 2 reports results for the baseline model when an alternative definition of savings is used: gross national savings rate. The results are consistent with those in column (1).

In column (1) of Table 3 we pool the data for sub sample 1 and sub sample 2 and introduce time and region dummies.[8] In Table 3, column (2) all explanatory variables are lagged in order to address possible endogeneity, as discussed above. The results are consistent with those in Table 2. Column (3) of Table 3 reports results when data are differenced over the two sub samples. This model eliminates any country specific fixed effects and serves as an alternative to the model in column (1) (where we explicitly control for fixed effects). It is interesting to note that both the level and squared term in column (3) are positive and significant. In Table 3, column (4) the model is estimated using panel data for all years and for all countries for which data are available. The fixed effects[9] model is preferred to the random effects model based on the Hausman test. The quadratic term is dropped as it is insignificant.[10] In columns (3) and (4) we see that the informal sector has a significant and positive impact on savings rates, *ceteris paribus*.

It is also possible to proxy the size of the informal sector with the share of self-employment in the total labor force, using data provided by the International Labor Organization. This is the approach favoured by Loayza and Rigolini (2011) because it is considered a better measure of the size of the informal sector especially in developing countries (see also Fiess et al. 2010). We estimated the baseline model with the share of self-employment measured in level form and as a quadratic. The latter was not significant so it was dropped and the model was re-estimated (column (5), Table 3). The result shows a positive relationship between the informal sector and savings rates.

The results for our other controls also warrant some discussion. There is no controversy concerning the relationship between income and savings. Tables 2 and 3 (columns (1)-(2)) clearly indicate that the relationship is non-linear; beyond some threshold level of income any further increase in income stimulates a fall in savings rates.[11] This result is consistent the observation that savings rates are lower in high income countries, where the emphasis is more on mass consumption (Ogaki et al. 1995; Masson et al. 1998, Freytab, 2013), and higher in low income countries where there is more precautionary saving (Schmidt-Hebbel and Serven; 1990; Smith, 2001; Chowdhury, 2015). The fact that low (high) income countries tend to have a large (small) informal sector means this result is also consistent with our a priori reasoning on savings and informality in Section 2. Growth is found to be inconsistent in its effect and was therefore omitted from some specifications.

The real rate of interest tends to be inversely related to savings rates (e.g., Loayza and Shankar, 2000; Ramajo et al. 2006; Swaleheen, 2008) but a positive link has been observed by Masson et al. (1998). This ambiguity reflects the relative strengths of the income and substitution effects arising from a change in the interest rate. Our results show a clear inverse relationship between the real interest rate

and savings rates. Trade openness tends to be positively related to economic performance (Harrison, 1996; Yanikkaya, 2003; Winters, 2004) and here a positive link between openness and savings is observed. Financial deepening is often measured by the money stock as a ratio of GDP or the flow of credit to the private sector. Loayza and Shankar (2000) show that private domestic credit exerted a negative impact on private savings rates, while the ratio of money to GDP had a statistically insignificant impact. Our results provide evidence to support the latter view.

The level of corruption, the dependency ratio, and inequality all have an adverse impact on savings rates. This is not surprising. Swalaheen (2008) discusses several channels - direct and indirect - through which corruption can erode savings. According to the life cycle hypothesis savings rates are lower when the population is very young or past retirement age. Inevitably, a higher dependency ratio leads to a drawing down on savings (Leff, 1964). With respect to inequality, one school of thought says higher inequality is associated with higher savings as the bulk of income lies with those who have a higher propensity to save. On the other hand, higher inequality may lead to lower income levels so savings rates are reduced (Barro, 2000).

The relationship between remittances and savings is an interesting result. A rise in remittances is associated with a fall in savings but in the gross savings model its impact is not significant. One would expect that a rise in remittances would boost disposable income and hence savings rates. However, if remittances are mainly devoted to consumption (Lipton, 1980) we would expect an inverse relationship. Macroeconomic stability (proxy using the inflation rate) is also frequently regarded as an important element in the savings decision. In our analysis the inflation rate is only significant in the model with lagged explanatory variables. Of the remaining variables included in at least one

specification, agriculture and the budget surplus/deficit ratio have a positive impact on savings rates while government consumption has a negative impact (Gale and Orszag, 2004; Loayza and Shankar, 2000).

5. Robustness of the non-linear result

One problem with estimating a non-linear relationship is that higher order terms place fairly strong restrictions on the link between savings and the informal sector that may not reflect the true underlying relationship. For instance, first and second order terms only allow for concave or convex relationships. In addition, the relationship has to hold over the entire data range. In contrast, semi-parametric methods allow possible non-linearity to be investigated in a more flexible manner while allowing for the (linear) effect of other control variables. The slope is allowed to vary across all sizes of the informal sector and is estimated at each point. Moreover, the estimator purges any non-linear correlation between the informal sector and other control variables. Though parametric and semi-parametric estimators provide different conclusions regarding non-linearity (e.g., Banerjee and Duflo, 2003) the semi-parametric results are likely to be more accurate.

We follow the semi-parametric methodology proposed by Robinson (1988) using the Kernel regression estimator.[12] Accordingly, one can consider the following equation to be estimated

$$S = \alpha + g(I) + \delta Z + u \qquad (2)$$

where Z is a set of explanatory variables that is assumed to have a linear effect on savings (S), $g(.)$ is a smooth and continuous, possibly non-linear, unknown function of I, and u is a random error term. Robinson's methodology proceeds in two steps.

In the first step $\hat{\delta}$ can be obtained by using OLS on

$$S - E(S|I) = \delta\,[Z - E(Z|I)\,] + v \qquad (3)$$

Where v satisfies $E(v|I,Z) = 0$ and $E(S|I)$ and $E(Z|I)$ are estimated using the Nadaraya-Watson non-parametric estimator. For instance, the estimation of $E(Y|I)$, $\hat{m}_S(I)$, can be written as

$$\hat{m}_S(I) = \underset{i}{\frac{\sum_{i=1}^{n} K_h(i - I_i) S_i}{\sum_{i=1}^{n} K_h(i - I_i)}} \qquad (4)$$

such that $i=1 \ldots n$ are the n number of observations, $K_h()$ is the shape function, commonly referred to as the Kernel. It is a continuous, bounded and real function that integrates to one and acts as a weighting function of observations around I and depends on the choice of bandwidth h. More specifically, this technique corresponds to estimating the regression function at a particular point by locally fitting constants to the data via weighted least squares, where those observations closer to the chosen point have more influence on the regression estimate than those further away, as determined by the choice of h and K. An additional appeal of this technique is that it avoids any parametric assumptions regarding $E(S|I)$ and thus about its functional form or error structure.

In the second step, the function g in (2) can be estimated by carrying out a non-parametric regression of $(S\text{-}Z)$ on I such that $\hat{\delta}$ is the OLS estimator of

$$S - \hat{m}_S(I) = \delta\big(Z - \hat{m}_Z(I)\big) + \varepsilon \qquad (5)$$

where ε is a random error term. Intuitively, $\hat{g}(I)$ is the estimate of $g(I)$ after the independent effect(s) of Z on S has been removed.[13] Given that the estimate of $\hat{g}(I)$ is at least in part based on non-

parametric estimation techniques, one cannot subject it to the standard statistical type tests, e.g., *t*-test. One can, however, relatively easily calculate upper and lower point-wise confidence bands as suggested by Härdle (1990).[14]

For all our estimations we use a Gaussian kernel for K_h and the optimal bandwidth *h* suggested by Fox (1990). One should note that the size of the estimated error variance, $\hat{\sigma}^2(I)$, at any point of *I* will depend proportionally on the marginal distribution of *I*. In other words, the accuracy of the estimate of *g(I)* at *I* is positively related to the density of other observations around that point. In order to visualize this effect, and as suggested by Härdle (1990), we calculate the point-wise confidence bands at points chosen according to the distribution of *I*. Specifically, we chose points so that one of the observations lie between them.[15] The explanatory control variables included when estimating (5) are those used in our parametric regressions.

The graph of $\hat{g}(I)$ along with confidence bands is shown in Figure 3. The relationship between the size of the informal sector and savings rates is not constant over the range of values for the informal sector. The relationship is negative when the informal sector is quite small and positive when the informal sector is larger. More specifically, the turning point occurs at just over 40, indicating that for countries with an informal sector that larger than 40 per cent of GDP there is a positive relationship between the informal sector and savings rates. The result confirms our earlier findings using regression analysis. What is interesting is that similar results are obtained using the panel data model as shown in Figure 4. This highlights the limited ability of regression models to capture non-linearity given that the linear fixed effects panel model suggested no non-linearity.

6. Conclusion

The relationship between savings rates and the informal sector is potentially very important in countries where the informal sector serves as a key source of employment and income. Yet until now little or nothing has been known about the empirical link between savings rates and the size of the informal sector. This paper has shown that there is a strong non-linear relationship using data for a large sample of countries. More specifically, when the informal sector is small there is an inverse relationship between savings rates and the size of the informal sector. However, beyond a certain point further increases in the size of the informal sector leads to a boost to savings rates. The U-shaped relationship between the size of the informal sector and savings is consistent with the idea that in countries with strong institutions (low levels of informality/later stage of development) savings fall because the indirect effect of the informal sector (lower savings due to lower output) outweighs the direct effect (higher savings for precautionary reasons) while in countries with weak institutions (high levels of informality/early stage of development) savings rise because the indirect effect (higher savings due to higher output) and the direct effect (higher savings for precautionary reasons) work in the same direction and reinforce one another. The non-linear effect is generally confirmed in both parametric and semi-parametric estimations.

Many countries require a boost to savings rates in order to stimulate further growth and development. The fact that savings rates are higher when the informal sector is well established (and growing) shows that informal firms are productive. So should we let the informal sector expand unchecked and mobilize savings this way? We do not believe this is a first-best solution because informal firms could be even more productive if they were able to operate in the formal sector. With formal status entrepreneurs will be able to access formal markets, make secure investments and gain new sources of

credit. The focus, therefore, should be on enabling informal firms to move successfully to the formal sector and remove barriers for business. This may be done through regulatory and tax reform as well as through anti-corruption measures aimed at reducing rent seeking behaviour by bureaucrats.

References

Alderman, H. (1996). Saving and economic shocks in rural Pakistan. *Journal of Development Economics*, 51: 343-365.

Andrade, R. and Guillen, A. (2014). Understanding domestic savings in Peru, Manuscript, Apoyo Consultoría.

Badaoui, E., Strobl, E., and Walsh, F. (2008). Is there an informal wage penalty: evidence from South Africa. *Economic Development and Cultural Change,* 56: 125-161.

Banerjee, A. and Duflo, E. (2003). Inequality and growth: what can the data say? *Journal of Economic Growth,* 3: 267-299.

Bargain, O. and Kwenda, P. (2010). Is informality bad: evidence from Brazil, Mexico and South Africa. Institute of Labour, Discussion Paper No. 4711.

Barro, R.J. (2000). Inequality and growth in a panel of countries. *Journal of Economic Growth,* 5: 5-32.

Beckmann, E., Hake, M., and Urvova, J. (2013). Determinants of households' savings in Central, Eastern and Southeastern Europe. The Oesterreichische Nationalbank. Vienna, Austria: Focus on European Economic Integration, Q3/13.

Bekkers, H. and Stoffers, W. (1995). Measuring informal sector employment in Pakistan: testing a new methodology. *International Labor Review,* 134: 17-36.

Benjamin, N. and Mbaye, A. (2012). The informal sector, productivity, and enforcement in West Africa: a firm-level analysis. *Review of Development Economics,* 16: 664-680.

Blundell, R. and Duncan, A. (1998). Kernel methods in empirical microeconomics. *Journal of Human Resources,* 33: 62-87.

Botero, J., Djankov, S., La Porta, R., Lopez-de-Silanes, F. and Shleifer, A. (2003). The Regulation of Labor. national Bureau of Economic Research. Working Paper no. 9756.

Browning, M. and Lusardi, A. (1996). Household saving: micro saving and micro facts. *Journal of Economic Literature,* 34: 1797-1855

Butelmann, A. and Gallego, F. (2000). Household saving in Chile: microeconomic evidence. *Journal Economía Chilena*, 3: 5-24.

Carroll, C. (1997). Buffer-stock saving and the Life Cycle/Permanent Income Hypothesis. *Quarterly Journal of Economics,* 112: 1-55

Carroll, C. (1998). Why do the rich save so much? National Bureau of Economic Research. Working Paper No. 6549.

Charmes, J. (2000). The contribution of the informal sector to GDP in developing countries: assessment, estimates, methods, orientations for the future. 4th meeting of the Delhi Group on Informal Sector Statistics, Geneva.

Choi, J.P. and Thum, M. (2005). Corruption and the shadow economy. *International Economic Review* 46: 817-836.

Chowdhury, A. (2015). Terms of trade shocks and private savings in the developing countries. *Journal of Comparative Economics*, 43: 1122-1134.

Cule, M. and Fulton, M. (2009). Business culture and tax evasion: why corruption and the unofficial economy can persist. *Journal of Economic Behaviour and Organisation*, 72: 811-822.

Dabla-Norris, E., Gradstein, M. and Inchauste, G. (2008). What causes firms to hide output? The determinants of informality. *Journal of Development Economics,* 85: 1-27.

De Soto, H. (1989). *The other path*. New York: Harper and Row.

Deaton, A. (1991). Saving and liquidity constraints. *Econometrica*, 59: 1221-1248.

Dell'Anno, R. and Schneider, F. (2003). The shadow economy of Italy and other OECD countries: what do we know? *Journal of Public Finance and Public Choice*, 21: 223-245.

Disney, R. (2006). Household saving rates and the design of public pension programmes: cross-country evidence. National Institute Economic Review, No. 198.

Djankov, S., La Porta, R., Lopez-de-Salinas, F., and Schleifer, A. (2002). The regulation of entry. *Quarterly Journal of Economics,* 117: 1-37.

Dobson, S. and Ramlogan-Dobson, C. (2012). Why is corruption less harmful to income inequality in Latin America? *World Development* 40: 1534-1545.

Docquier, R., Muller, T. and Naval, J. (2017). Informality and long run growth. *Scandinavian Journal of Economics*, 119: 1040-85.

Dreher, A., Kotsogiannis, C. and McCorriston, S. (2009). How do institutions affect corruption and the shadow economy? *International Tax Public Finance* 16: 773–796

Elgin, C. and Uras, R.B. (2013). Is informality a barrier for financial development? *SERIEs* 4: 309-331.

Feige, E.L. (1989). *The Underground Economies: Tax Evasion and Information Distortion*. Cambridge: Cambridge University Press.

Feige, E.L. (1994). The underground economy and the currency enigma. *Supplement to Public Finance/ Finances Publiques*, 49: 119-136.

Feiss, N., Fugazza, H., and Maloney, W.F. (2010). Informal self-employment and macroeconomic fluctuations. *Journal of Development Economics* 91: 211-226.

Fox J. (1990). Describing univariate distributions, in Fox J. and Long J.S. (eds), *Modern methods of data analysis*, Sage Publications.

Freytab, A. (2013). Institutions and savings in developing and emerging economies. *Public Choice*, 157: 475-509

Friedman, M. (1953). *Essays in Positive Economics*. Chicago: Chicago University Press.

Gale, W.G. and Orszag, P. (2004). Budget deficits, national savings and interest rate. Paper prepared for the Brookings Panel on Economic Activity, September.

Gelb, A., Mengistae, T., Ramachandran, V., and Shah, M.K. (2009). To formalize or not to formalize? Comparisons of microenterprise data from Southern and East Africa. Available at: http://siteresources.worldbank.org/INTMACRO/Resources/InformalFirmsAfricaGelbMengistaeRamachandran-Shah1292009.pdf.

Granada, C. and Hamann, F. (2015). Informality, saving and wealth inequality. *Borradores de Economia.* No. 873.

Grigoli,F., Herman, A. and Schmidt-Hebbel, K. (2014). World saving. International Monetary Fund Working Paper. WP/14/204.

Härdle, W. (1990). Applied non-parametric regression. Econometric Society Monographs Series 19, Cambridge University Press.

Harrison, H. (1996). Openness and growth: a time series, cross-country analysis for developing countries. *Journal of Development Economics*, 48: 419-447.

Horioka, C.Y., and Hagiwara, A. (2011). The determinants and long term projections of savings rates in developing Asia. National Institute Economic Review, Working Paper No. 17581.

Johnson, J., Kaufman, D. and Shleifer, A. (1997). The unofficial economy in transition. *Brookings Papers on Economic Activity*, 2: 159-239

Johnson, S., Kaufmann, D., Shleifer, A., Goldman, M. I., & Weitzman, M. L. (1997). The unofficial economy in transition. *Brookings Papers on Economic Activity*, 2, 159–239.

Khamis, M. (2012) Is informal sector work an alternative to workfare benefits? The case of pre-program expansion and economic crisis. *Review of Development Economics,* 16: 579-593.

Leff, N. H. (1964). Dependency rates and savings rates. *American Economic Review* 68: 886-898.

Lipton, M. (1980). Migration of rural areas of poor countries: the impact on rural productivity and income distribution. *World Development*, 8: 1-24.

Loayza, N. (1996). The economics of the informal sector: A simple model and some evidence from Latin America. *Carnegie-Rochester Conference Series on Public Policy*, 45: 129-162.

Loayza, N. (2016). Informality in the process of development and growth, Policy Research Working Paper Series 7858, The World Bank.

Loayza, N., and Rigolini, J. (2011). Informal employment: safety net or growth engine. *World Development,* 39: 1503-1515.

Loayza, N., and Shankar, R. (2000). Private savings in India. *The World Bank Economic Review,* 14: 571-94.

Loayza, N., Oviedo, A.M., and Serven, L. (2005). The impact of regulation on growth and informality. *World Bank Staff Papers,* 3623.

Loayza, N., Schmidt-Hebbel, K. and Serven L. (2000). What drives private saving across the world. *Review of Economic and Statistics,* 82: 165-181.

Masson, P., Bayoumi, T and Samiei, H. (1998). International evidence on the determinants of private savings. *World Bank Economic Review,* 12: 483-501.

McKenzie, D. and Sakho, Y. (2010). Does it pay firms to register for taxes? The impact of formality on firm profitability, *Journal of Development Economics,* 91: 15-34.

Modigliani, F., and Brumberg, R. (1954). Utility analysis and the consumption function: an interpretation of cross-section data, in Kurihara, K. (ed.), *Post-Keynesian Economics*, Rutgers University Press, New Brunswick.

Ogaki, M., Ostry, J. and Rheinart, C. (1995). Savings behavior in low and middle income countries: a comparison. *International Monetary Fund Working Paper,* WP/95/3.

Ordonez, J. (2014). Tax collection, the informal sector, and productivity. *Review of Economic Dynamics*, 17: 262-86.

Paxson, C.H. (1992). Using weather variability to estimate the response of savings to transitory income in Thailand. *American Economic Review*, 82: 15-33.

Ramajo, J., Garcia, A. and Ferre, M. (2006). Explaining aggregate savings behaviour: new evidence from a panel of OECD countries. *Applied Financial Economics Letters,* 2: 311-315.

Robinson, P. (1988). Root-N-consistent semiparametric regression. *Econometrica*, 56: 931-954.

Sarte, P.D.G. (2000). Informality and rent seeking in a model of long run growth. *Journal of Monetary Economics*, 46: 173-197.

Schmidt-Hebbel, K. and Serven, L. (1999). Introduction. In *The Economics of Saving and Growth*, edited by K. Schmidt-Hebbel and L. Servén. Cambridge University Press.

Schneider, F. (2012). The shadow economy and work in the shadow: what do we (not) know? Institute for the Study of Labor Discussion Paper No. 6423.

Schneider, F., Buehn, A. and Montenegro, C. (2010). New estimates for the shadow economies all over the World. *International Economic Journal*, 24: 443-461.

Smith, D. (2001). International evidence on how income inequality and credit market imperfections affect private saving rates. *Journal of Development Economics*, 64: 103-127.

Smith, P. (1994), Assessing the size of the underground economy: the Canadian statistical perspectives. *Canadian Economic Observer*, Catalogue No. 11-010: 16-33.

Swaleheen, M. (2008). Corruption and savings in a panel of countries. *Journal of Macroeconomics*, 30: 1285-1301.

Tanzi, V. (1999), Uses and abuses of estimates of the underground economy. *Economic Journal*, 109/3: 338-347.

Tsikata, Y., (1998). Saving in South Africa: An empirical investigation. (unpublished; Washington: World Bank).

Vergara, R. (2001). Determinantes del ahorro privado," in F. Morandé, R. Vergara, N. Loayza, and K. Schmidt-Hebbel (eds.): Análisis Empírico del Ahorro en Chile, Vol. 1, Central Bank of Chile, Santiago, Chile, pp. 83-104.

Verick, S. (2008). The impact of globalization on the informal sector in Africa. Economic and Social Policy Division, UNECA, Addis Ababa.

Wand, M.P. and Jones, M.C. (1995). *Kernel smoothing*. Chapman & Hall, London.

Winters, L.A. (2004). Trade liberalization and economic performance: on overview. *Economic Journal,* 114: F4-F21.

Yanikkaya, H. (2003). Trade openness and economic growth: a cross-country empirical analysis. *Journal of Development Economics* 72: 57-89.

Notes

[1] Empirical studies show that the key drivers of informality are high taxes, a lack of social security and benefits, excessive regulation, culture, and poor governance (e.g., Loayza, 1996; Johnson et al. 1997; Tanzi, 1999; Botero et al. 2003; Dabla-Norris et al. 2008; Khamis, 2012; Ordonez, 2014).

[2] To test these theories one needs to disaggregate income into its permanent and transitory components (e.g., Paxson, 1992; Alderman 1996; Loayza and Shankar, 2000; Vergara, 2001; Andrade and Guillen, 2014; Grigoli et al. 2014). The empirical results are inconclusive. Testing these theories is beyond the scope of this paper.

[3] Research has shown that households may save more when facing periods of low income (e.g., Deaton, 1991; Carroll, 1998; Butellman and Gallego, 2000).

[4] In essence, in countries with weak institutions (early stage of development) the choice is between no production and production in the informal sector, while in countries with strong institutions (late stage of development) the choice is between production in the formal sector and the informal sector. This helps to explain why the informal sector makes a positive (negative) contribution to output in countries with weak (strong) institutions.

[5] Available at http://www.wider.unu.edu/wiid/wiid.htm.

[6] Available at http://pwt.econ.upenn.edu/php_site/pwt_index.php.

[7] Savings = -113.8559 - 0.1104informal +31.1805income -1.6066incomesq +35.1726growth -1.3859remittances
(0.69) (0.216) (0.048) (0.104) (0.466) (0.000)

-1.5351corrupt +0.0358open -0.2973interest -0.2498inflation
(0.128) (0.090) (0.019) (0.027)

R^2=0.57 n=122

[8] Results available upon request.

[9] It is not possible to estimate a dynamic version of the model because of missing data and because informal sector data are only available for a short time period.

[10] Results available upon request.

[11] The U-shaped relationship between savings and income holds if GNP is used rather than GDP:

Savings = -13.2349 –0.7469 informal + 0.0082informalsq+10.8366income-0.5269incomesq-0.0729dscp
(0.489) (0.015) (0.029) (0.017) (0.055) (0.001)

+0.0438open -0.2907interest
(0.000) (0.001)

[12] See Blundell and Duncan (1998) for details and a helpful discussion of the implementation of this method.

[13] Note that this also involves controlling for non-linear effects between Z and I.

[14] It is worth noting that the confidence band proposed by Härdle (1990) ignores possible approximation error bias. Correcting for this would complicate the expression considerably since the bias is a complicated function of the first and second derivatives of g(I). The bias tends to be highest at sudden peaks of the data and at the necessarily truncated left and right boundaries of the data. However, if h is chosen proportional to 1/n(1/5) times a sequence that tends slowly to zero then the bias vanishes asymptotically for the interior points; see Härdle (1990) and Wand and Jones (1995).

[15] For the endpoints we chose the 1 and 99 per centiles of the distribution.

[16]Alternatively we could start with a production function of the form $Y = K^{\alpha}(CAL)^{1-\alpha}$ (where C refers to rent seeking activities) and work through to the same steady state solution.

[17] A similar equation is in Swalaheen (2008) but no explanation is given as to its derivation.

Table 1. *Summary statistics*

Variable	Mean	Std Dev	Max	Min	Obs.
Informal	34.58	13.83	68.30	8.40	161
Gross National Savings	20.70	11.75	98.25	-7.31	121
Gross Domestic Savings	17.60	16.96	81.29	-29.95	178
GDP	11177.57	12993.60	63317.45	202.08	189
Openness	89.61	47.73	383.08	1.84	189
Gini Coefficient	41.39	8.97	65.377	21.9	130
Corruption (CPI)	4.02	2.07	9.71	1.44	179
Corruption (ICRG)	2.63	1.10	6.00	0.33	139
Dependency Ratio	64.31	18.04	108.59	29.02	181
Inflation	25.40	229.44	2983.90	-1.4326	169
Interest Rate	8.20	17.01	203.44	-16.69	162
Money to GDP Ratio	57.49	59.29	610.36	5.74	180
Domestic Credit to Private Sector	46.85	44.28	215.119	1.72	181
Budget Surplus/Deficit	-0.76	3.66	13.88	-12.21	136
Political Stability	-0.14	0.92	1.64	-2.35	141
Remittances	3.70	5.34	29.89	0.018	188
Agriculture	15.24	13.78	68.29	0.07	180
Government Consumption	18.68	8.07	83.16	3.79	174
Growth	0.036	0.035	0.281	-0.051	189

Table 2. *Savings estimations: base models*

Explanatory Variables	Dependent variable: Gross domestic savings	Gross national savings

	(1) (entire sample)	(2) (sub sample 1)	(3) (sub sample 2)	(4) (entire sample)
Informal	-1.0489*** (0.003)	-0.6068** (0.042)	-1.3099*** (0.006)	-0.9416*** (0.001)
Informal^2	0.0125*** (0.006)	0.0061* (0.073)	0.0162*** (0.009)	0.0099*** (0.006)
Income	33.8084** (0.024)	24.3185** (0.024)	46.1989*** (0.001)	21.5277* (0.0813)
Income^2	-1.7848** (0.049)	-1.3510* (0.053)	-2.4000** (0.031)	-1.3245* (0.082)
Growth	24.0664 (0.587)			44.7074 (0.223)
Remittances	-1.3894*** (0.000)	-1.0114*** (0.000)	-1.3547*** (0.000)	-0.2916** (0.043)
Corruption	-2.0146* (0.057)	-0.3688 (0.610)	-2.4901** (0.040)	-1.0442 (0.229)
Openness	0.0412* (0.053)	0.0518*** (0.008)	0.0320 (0.111)	0.0390*** (0.007)
Interest rate	-0.3056** (0.025)	-0.2130** (0.019)	-0.2625* (0.057)	-0.2076* (0.076)
Inflation	-0.2119** (0.048)			-0.1408* (0.093)
Surplus/deficit				0.8434*** (0.004)
Constant	-105.9709* (0.086)	-74.1093* (0.076)	-160.2916** (0.042)	-40.0813 (0.538)
R^2	0.60	0.55	0.61	0.41
Observations	122	91	120	99

P values in parentheses are based on heteroskedastic-robust standard errors.
*significant at 10%; ** significant at 5%; *** significant at 1%.
Differences in sample size reflect model specification and data availability (baseline model, models for two sub samples, and model with alternative definition of savings).

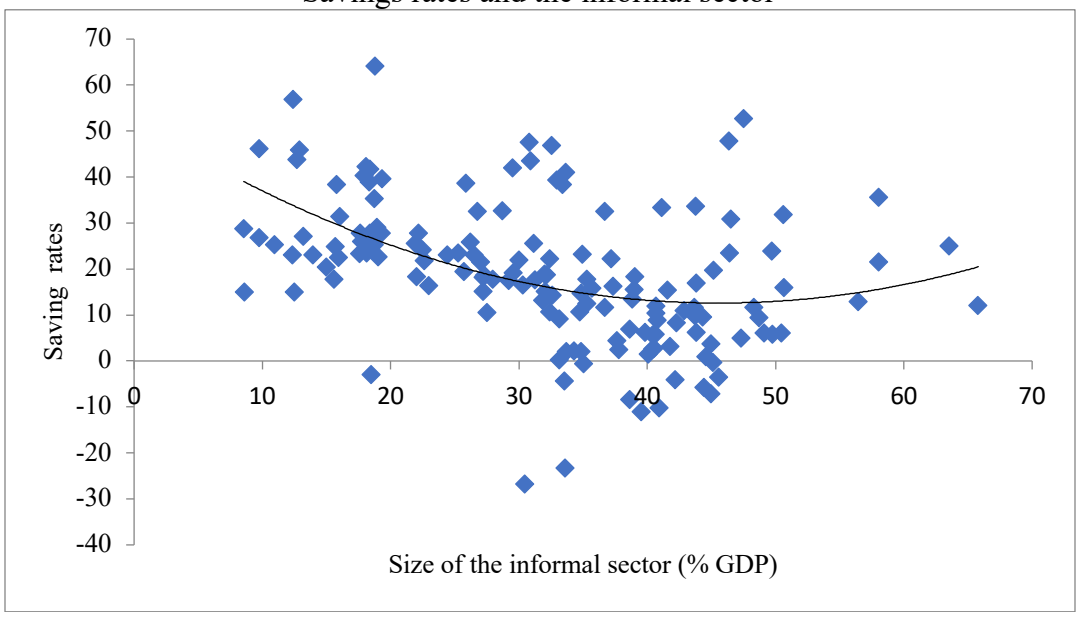
Table 3. *Savings estimations: alternative models*

Explanatory Variables	Dependent variable: Gross domestic savings				
	(1)$^{\wedge}$	(2)*	(3)$^{+}$	(4)$^{\&}$	(5)$^{\$}$

Informal	-1.0593*** (0.000)	-0.8484*** (0.009)	5.0650*** (0.007)	0.8388** (0.031)	12.0.6554* (0.075)
Informal^2	0.01399*** (0.000)	0.0107** (0.016)	1.2533** (0.025)		
Income	34.6824** (0.012)	49.5311*** (0.001)	21.096*** (0.000)	18.7018*** (0.000)	38.5698** (0.015)
Income^2	-1.4729 (0.077)	-2.9092*** (0.002)			-1.9661** (0.028)
Growth					104.534** (0.024)
Remittances	-1.0126*** (0.000)	-1.6529*** (0.000)	-0.7592** (0.013)	-0.6735**** (0.000)	-1.2496*** (0.000)
Corruption	-1.8271*** (0.001)	-0.0942 (0.901)	-1.4253 (0.215)	-0.7535* (0.084)	
Openness	0.03230** (0.016)	0.0423** (0.043)			
Interest rate	-0.1815*** (0.003)	-0.2246** (0.047)	-0.2066** (0.021)	-0.0664*** (0.003)	
Inflation		-0.0625 (0.0304)			
Agriculture			0.1497 (0.629)	0.04390** (0.024)	
Dependency			-0.4195* (0.070)		0.1315 (0.336)
Credit	-0.0750*** (0.002)		-0.0491** (0.046)		
Gini coefficient				-0.177* (0.071)	
Government consumption					-0.0649*** (0.001)
Broad money				-0.0440** (0.016)	0.0650** (0.046)
Constant	-140.8944** (0.032)	-164.3121*** (0.004)	0.5578 (0.610)	-168.7865*** (0.000)	-168.5552** (0.027)
R^2	0.70	0.55	0.48	0.46	0.65
Observations	211	89	81	309	77

p values in parentheses are based on heteroskedastic-robust standard errors.
*significant at 10%; ** significant at 5%; *** significant at 1%.
Differences in sample size reflect model specification and data availability: $^{\wedge}$ estimation over two sub periods with region and time dummies; * estimation over full sample period with lagged explanatory variables; $^{+}$ estimation over two sub periods and differenced data; $^{\&}$ estimation over full sample period with panel data; $^{\$}$ estimation over full sample period with share of self-employment as measure of size of the informal sector.

Figure 1
Savings rates and the informal sector



Note: The U-shape is fitted in Excel using the trend-line option; this fits a polynomial regression of order two (a best fit quadratic line).

Figure 2
Savings rates and the informal sector (>38% of GDP)

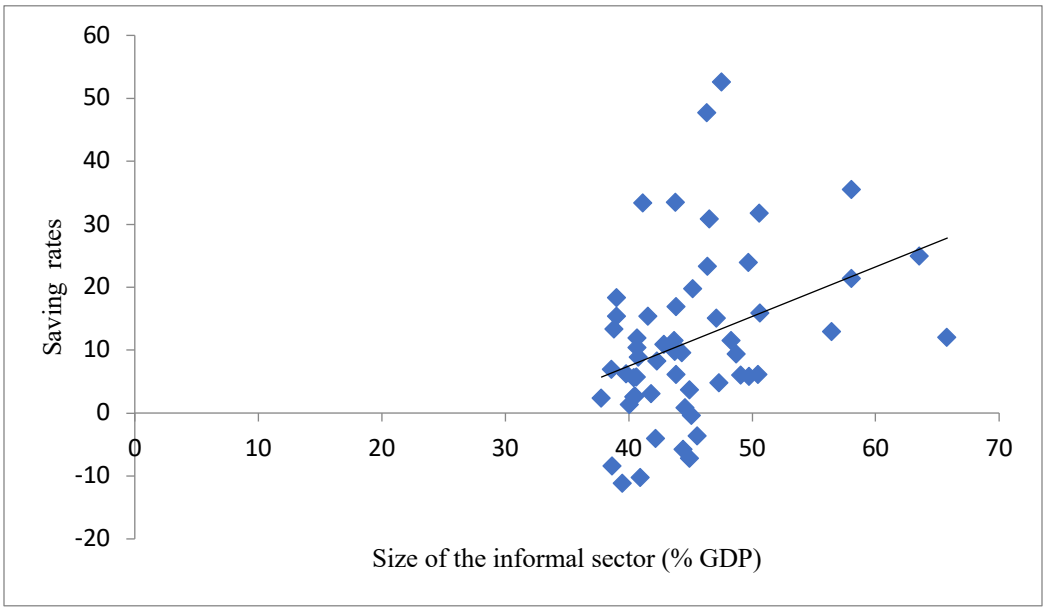


Figure 3
Savings and informality: semi-parametric estimate

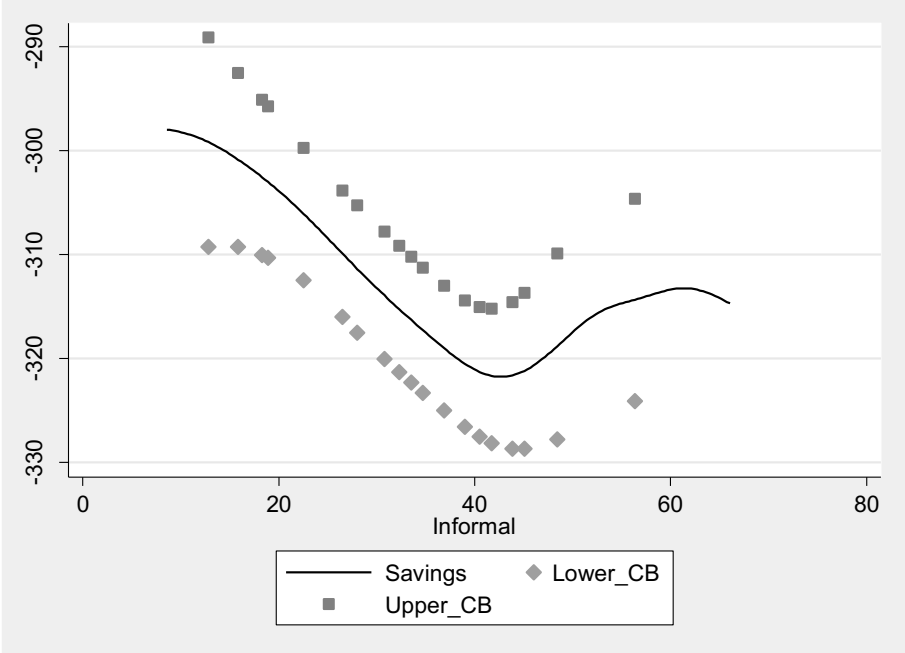
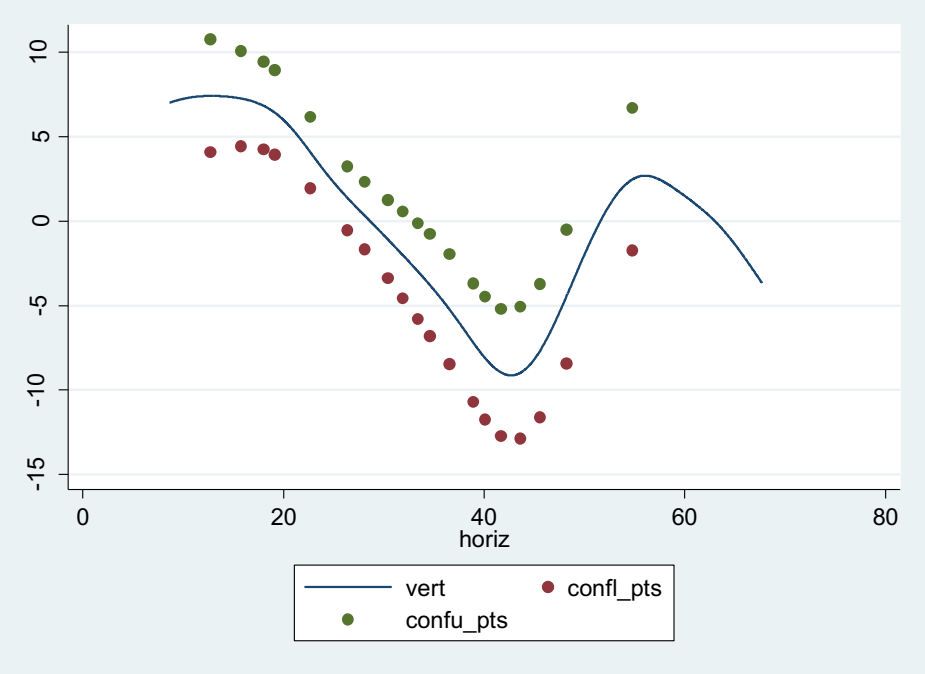


Figure 4
Savings and informality: semi-parametric estimate (panel model)



Appendix

To examine the link between the informal sector and output consider a typical Cobb-Douglas production function

$$Y = K^{\alpha}(AL)^{1-\alpha} \quad \text{(A1)}$$

Y is output, K is capital, L is labour and A is a parameter to represent the efficiency or effectiveness of the workforce. In addition to standard assumptions we assume that rent-seeking costs (illegal payments made by firms to start and maintain continuous operation in the formal sector) are subsumed in A. Rewriting the production function in terms of output per effective worker (dividing by AL) gives

$$y = k^{\alpha} \quad \text{(A1a)}$$

where $y = \frac{Y}{AL}$ and $k = \frac{K}{AL}$

Taking the total derivative of k

$$dk = \frac{\partial k}{\partial K}dK + \frac{\partial k}{\partial L}dL + \frac{\partial k}{\partial A}dA = \frac{dK}{AL} - k\frac{dL}{AL} - k\frac{dA}{AL} \quad \text{(A2)}$$

And given that $\frac{dK}{AL} = sk^{\alpha} - \delta k; \frac{dL}{L} = n; \frac{dA}{A} = g$

(A2) may be re-written as

$$dk = sf(k) - (\delta + n + g)k \quad \text{(A3)}$$

s is the savings rate, δ is the rate of depreciation, n is the population growth rate and g is the growth rate of efficiency or worker effectiveness. Since A includes rent seeking costs, g can be broken down to include *a* and *r*, where a refers to technical progress and r refers to rent seeking,[16] so that

$$dk = sf(k) - (\delta + n + a + r)k \quad \text{(A4)}[17]$$

The rent seeking element is likely to be substantial and not re-invested (Braun and Loayza, 1993). For example, Djankov et al. (2002) find that in more than one third of the countries in their sample, rent seeking costs accounted for more than 50 per cent of GDP per capita. In the case of the Dominican Republic, such costs are in excess of four and a half times GDP per capita.

Following from (A4), in steady state

$$sf(k^*) = (\delta + n + a + r)k^* \tag{A5}$$

Also, since we know $f(k) = k^{\alpha}$ the model can be solved to give steady state levels of k and y

$$k^* = \left[\frac{s}{(\delta + n + a + r)}\right]^{\frac{1}{1-\alpha}} \tag{A6}$$

and

$$y^* = \left(\frac{s}{(\delta + n + a + r)}\right)^{\frac{\alpha}{1-\alpha}} \tag{A7}$$

Equations (A6) and (A7) make clear the link between rent seeking activities, capital per person and output per person; as firms choose to operate in the informal sector, rent-seeking (r) by corrupt officials is reduced and both k^* and y^* rise. In effect, the informal sector substitutes for production in the formal sector and total output rises. We may also interpret *r* to include costs associated with regulation and bureaucratic requirements. Djkanov et al. (2002) points out that the principal beneficiaries of regulation are bureaucrats and politicians so to the extent that growth of the informal sector reduces this cost, output will rise.

Table A1. *Variable definition and source*

Variable	Definition	Source

Informal	Size of the informal sector measured as percentage of official GDP	Schneider et al. (2010) (see also, Schneider, 2012)
Gross National Savings	Gross national disposable income less final consumption expenditure	World Development Indicators (2011)
Gross Domestic Savings	Gross domestic disposable income less final consumption expenditure	World Development Indicators (2011)
GDP	Real GDP per capita	Penn World Tables 7 http://pwt.econ.upenn.edu/php_site/pwt_index.php.
GNP	Real GNP per capita	World Development Indicators (2017)
Openness	Exports plus imports as a ratio of GDP	Penn World Tables 7 http://pwt.econ.upenn.edu/php_site/pwt_index.php.
Gini Coefficient	Measure of inequality using the Lorenz curve	WIDER http://www.wider.unu.edu/research/Database/en_GB/wiid/
Corruption (CPI)	Abuse of entrusted power for private gain	CPI index http://www.transparency.org/research/cpi/overview
Corruption (ICRG)	An assessment of corruption within the political system	ICRG index http://www.prsgroup.com/ICRG.aspx
Dependency Ratio	Ratio of dependents - people younger than 15 or older than 64 - to the working-age population	World Development Indicators (2011)
Inflation	Annual growth rate of GDP implicit deflator	World Development Indicators (2011)
Interest Rate	Rate paid by commercial banks for demand, time, or savings deposits	World Development Indicators (2011)
Money to GDP Ratio	Sum of currency outside banks, demand deposits, time, savings, and foreign currency deposits	World Development Indicators (2011)
Domestic Credit to Private Sector	Financial resources provided to the private sector as percentage of GDP	World Development Indicators (2011)
Budget Surplus/Deficit	Revenue (including grants) minus expenses, minus net acquisition of nonfinancial assets	World Development Indicators (2011)
Political Stability	Difference between autocracy and democracy scores	Polity Data Base http://www.systemicpeace.org/polity/polity4.htm
Remittances	Sum of transfers by migrant workers and wages and salaries earned by non resident workers	World Development Indicators (2011)
Agriculture	Forestry, hunting, fishing, as well as cultivation of crops and livestock production	World Development Indicators (2011)

	(value added as percentage of GDP)	
Government Consumption	All government current expenditures for purchases of goods and services	World Development Indicators (2011)
Growth	Rate of increase of real GDP per capita	Penn World Tables 7 http://pwt.econ.upenn.edu/php_site/pwt_index.php.