Tax evasion, financial development and inflation: theory and empirical evidence

Manoel Bittencourt[*]   Rangan Gupta[†]   Lardo Stander[‡]

February 2013

Abstract

Using a standard overlapping generations monetary production economy, faced with endogenously determined tax evasion by heterogeneous agents in the economy, we provide a theoretical model that indicates that both a lower (higher) level of financial development and a higher (lower) level of inflation leads to a bigger (smaller) shadow economy. These findings are empirically tested within a panel econometric framework, using data collected for 150 countries over the period $1980-2009$ to enable a broad generalisation of the results. The results support the developed theoretical model, even after having accounted for the differences in the levels of economic development, the level of institutional quality that includes different tax regimes and regulatory frameworks, central bank participation in the economy as well as different macroeconomic policies.

JEL Classification: C61, E26, P16
Keywords: Informal economy, financial development, inflation.

---

[*]Department of Economics, University of Pretoria, Pretoria, 0002, South Africa.
[†]To whom correspondence should be addressed. Department of Economics, University of Pretoria, Pretoria, 0002, South Africa, Email: rangan.gupta@up.ac.za.
[‡]Department of Economics, University of Pretoria, Pretoria, 0002, South Africa.

1 Introduction

"To think of shadows is a serious thing." - Victor Hugo[1]

Recent empirical evidence provided by Bose, Capasso and Wurm (2012) show that an improvement in the development of the banking sector is associated with a smaller shadow economy. The findings of Bose et al. (2012) corroborate indicative theoretical results reported by Blackburn, Bose and Capasso (2010) that a less-developed financial sector corresponds to the observance of a bigger shadow economy. Blackburn et al. (2010) studied the relationship between the underground economy and financial development in a model of tax evasion and bank intermediation. In their model, agents with heterogeneous skills seek loans in order to undertake risky investment projects, with asymmetric information between borrowers and lenders implying a menu of loan contracts that induce self-selection in a separating equilibrium. Given these contracts, agents choose how much of their income to declare by trading off their incentives to offer collateral against their disincentives to comply with tax obligations. The main implication of the analysis is that the marginal net benefit of income disclosure increases with the level of financial development. Thus, as with the empirical observation made by Bose et al. (2012), the paper shows that the lower is the stage of such development, the higher is the incidence of tax evasion and the greater is the size of the underground economy. Furthermore, Gupta and Ziramba (2010) using an overlapping generations (OLG) monetary endogenous growth model, whereby government transfers affect young-age income, show that inflation - besides the usual suspects like fiscal policy (Dabla-Norris and Feltenstein, 2005), penaly rates (Schneider, 1994), probability of being detected (Schneider and Enste, 2000) and degree of corruption (Cerqueti and Coppier, 2011) - affect the degree of tax evasion. Specifically, they indicate a negative relationship between inflation and the fraction of income reported.

Against this backdrop, the objectives of this paper are twofold: First, using a monetary OLG stochastic production economy, characterised by endogenous tax evasion, we provide a novel theoretical explanation that both lower financial sector development as well as higher inflation (money growth rate) leads to a bigger shadow economy, and; second, with the theoretical analysis presented yielding an empirically-testable equation (albeit not in the sense of a one-to-one correspondence) relating tax evasion with financial development and inflation, we test the validity of the theoretical implications using a panel of 150 countries for the period 1980 to 2009, based on a newly-constructed dataset of shadow economy estimates by Elgin and Öztunali (2012).[2] To the best of our knowledge, this paper is not only the first

---

[1] *Les Misérables* (1862).

[2] Note that the shadow economy estimates of Elgin and Öztunali (2012) is obtained

attempt at providing a simultaneous theoretical explanation of how both (lower) financial development and (higher) inflation may lead to (higher) tax evasion and therefore, to the observance of a (bigger) shadow economy,[3] but also empirically corroborate the theoretical claims.

At this stage, it is important to put into context the importance of our theoretical result that monetary policies (money growth rate and cash-reserve requirements held by financial intermediaries[4] could also affect the level of tax evasion. Gupta (2008) and Gupta and Ziramba (2009) point out that studies (such as Roubini and Sala-i-Martin (1995), Gupta (2005) and Holman and Neanidis (2006)) which analyse optimal (growth- and/or welfare-maximising) mix of fiscal and monetary policy suffer from the Lucas (1976) critique, by treating tax evasion exogenously. Gupta (2008) and Gupta and Ziramba (2009) reached such conclusions by developing growth models with tax evasion being a behavioural decision (as also pointed out theoretically by Chen (2003), Arana (2004) and Atolia (2007)) to indicate that the level of tax evasion is dependant on the tax and penalty rates. Given this, following a change in the degree of tax evasion, the tax and the penalty rates are not available to the policy maker to respond optimally to such a change, since clearly changes in these policy variables would affect the level of tax evasion further. Thus, Gupta (2008) and Gupta and Ziramba (2009) studies optimal monetary policy response following changes in the degree of tax evasion emanating from not only movements in the structural parameters of the model, but also variations in the tax and penalty rates.[5] Now, with tax evasion also affected by monetary policy, it would imply that the studies of Gupta (2008) and Gupta and Ziramba (2009) is not immune to the Lucas (1976) critique either. In summary, studies that analyse optimal (growth- and/or welfare-maximising) monetary and fiscal policy following a change in the degree of tax evasion is likely to lead to non-optimal policy outcomes, since changes in the policy parameters in response to the change in the level

from a calibrated dynamic general equilibrium model for various countries over different periods.

[3]We concede that tax evasion and shadow economy are not necessarily synonymous, but contend that measures of the shadow economy are systematically used in the literature as a proxy for the level of tax evasion (Alm, 2012). The use of tax evasion as a substitute for the shadow economy also resonates with the adopted definition of the shadow economy in this paper, and facilitates the theoretical approach followed. Moreover, following Gupta (2005) it can be shown that $\frac{TE}{Y} = SE * \tau$, where $\frac{TE}{Y}$ is tax evasion as a percentage of gross domestic product (GDP), SE is a measure of the shadow economy and τ is a parameter measuring taxes paid as a percentage of GDP.

[4]Note that, the cash-reserve requirements have been long viewed as a measure of financial repression, since higher the cash reserve requirements, lesser the loans available to a bank to lend out for investment/production purposes. For a detailed discussion along these lines, refer to Gupta (2005, 2008) and Gupta and Ziramba (2009, 2010).

[5]See Koreshkova (2006) for a similar analysis relating inflation and the underground economy, where the shadow economy is modelled by distinguishing between a formal and informal production structure, instead of endogenous tax evasion.

of tax evasion (arising from changes in the structural parameters affecting the degree of evasion) would change the degree of tax evasion further.

The rest of the paper is organised as follows: Section 2 describes the economic setting for our analysis; Sections 3 - 5, respectively, defines the competitive equilibrium, solves the model for the optimal degree of the shadow economy, discusses the empirical evidence obtained from our dataset against the current background to the observance of the shadow economy and Section 5 offers some concluding remarks.

2 The economic setting

Time is divided into discrete segments and indexed by $t = 1, 2, ...$ The principal economic activities are: (i) entrepreneurs who live for two periods, receive a positive young-age endowment of W_1 and consume only when old. When the cost of undertaking an investment project exceeds the current endowment of entrepreneurs, they require external finance. To obtain the external finance, entrepreneurs have to offer collateral to the banks and thus have to decide what portion of their income to declare in order to increase the probability of obtaining external finance. This external finance is provided by the banks according to the terms and conditions of optimal loan contracts; (ii) each two-period lived overlapping generations depositor receives a young-age endowment of $0 \leq W_2 \leq 1$ and an old-age endowment of $0 \leq W_3 \leq 1$. The depositors consume in both periods. The young-age consumer evades a portion of the tax-liability, with the tax evasion being determined endogenously to maximize utility, and the remainder is allocated either towards young-age consumption or deposited in the banks, for future old-age consumption; (iii) the banks operate in a competitive environment and perform a pooling function by collecting the deposits from the consumers and lending it out to the entrepreneurs after meeting an obligatory cash reserve requirements; and (iv) there is an infinitely-lived consolidated government which meets its non-productive expenditure by taxing income, generating seigniorage income and setting a penalty for tax evasion when caught. The government also controls its two main policy instruments, namely money growth rate and the reserve requirement. The government balances its budget on a period-by-period basis. There is a continuum of each type of economic agent with unit mass.

We introduce ex-post moral hazard into the economy due to banks facing a costly state verification (CSV) problem since entrepreneurs can declare bankruptcy even when they are not. The principal outcome of those investment projects of the entrepreneurs, financed via bank loans, is essentially private information to the entrepreneur. If banks are willing to incur some monitoring cost, they can observe the same outcome. Note that the size of CSV is used here as a "proxy" for the efficiency of the financial system. In

4

line with Di Giorgio (1999) and Gupta (2005), it is reasonable to assume that a more developed financial system will have a lower CSV.

2.1 Entrepreneurs

Entrepreneurs live for two periods, receive an initial endowment of W_1, undertake some type of investment and only consumes in the second period. They have access to a simple investment technology such that by investing one unit of the consumption good at t, either $\alpha > 1$ units are produced at $t+1$ with probability of q or 0 units are produced with probability of $1-q$. Capital investment undertaken by the entrepreneur, K_t, is limited by the availability of funding to the entrepreneurs. Hence:

$$K_t = W_1 + l_t \tag{1}$$

where $l_t = \frac{L_t}{p_t}$ and L_t is the nominal quantity of loans that entrepreneurs can obtain from the banks. If the investment activity of the entrepreneur is successful, the cost of external finance obtained at time point t that is repaid to the bank, is a gross interest rate of $1 + i_{lt+1}$. If the investment activity is not successful, resulting in the entrepreneur declaring bankruptcy, nothing is repaid to the bank. The level of output produced by the entrepreneur at time point $t+1$ with probability q, is then:

$$y_{t+1} = \alpha K_t \tag{2}$$

or 0 with probability $1-q$. Thus, the entrepreneur's consumption in the second period, C_{t+1}^e depends on the initial endowment of W_1; the yield of the investment, α; the cost of the external finance obtained from the banks, $1 + i_{lt+1}$ and the probability of success, q. Taking $1 + \pi_{t+1} = \frac{p_{t+1}}{p_t}$, the gross inflation rate and replacing (1) into (2), the entrepreneur's problem is precisely defined as:

$$C_{t+1}^e = \alpha(W_1 + l_t) - (1 + i_{lt+1})\frac{l_t}{1 + \pi_{t+1}} \tag{3}$$

with probability of q or

$$C_{t+1}^e = 0 \tag{4}$$

with probability of $1-q$. As the outcome of the entrepreneur's problem is intertwined with the outcome of the bank's problem, the problem will not be explicitly solved here but rather as part of the bank's problem.

2.2 Depositors

All depositors have the same preferences, so there is a representative agent in each period. Depositors receive an initial young-age endowment of W_2

and an old-age endowment of W_3, respectively. Both age-type endowments obey $0 \leq W_2, W_3 \leq 1$, and it follows that $\sum_{i=1}^{3} W_i = 1$. Thus, at time point t, there are two coexisting generations of young-age and old-age depositors. N people are born at each time point $t = 1$. At time point $t = 1$, there exist N people in the economy called the initial old, who live for only one period and at each time point $t = 1$, N people are born (the young generation) and N people are beginning the second period of their life (the old generation). Note, the population N here is assumed to be constant, therefore N is normalized to 1.

The government sets a tax of rate τ on the young-age endowment received by the depositor, which can be evaded - at a cost[6] - with a given probability of σ. Thus, for the potential evader, there exists the possibility of two tax states: 'success' (getting away with evasion) or 'failure' (being discovered and incurring a penalty) with the probability of $1 - \sigma$. The depositor knows *ex-ante* the probability of getting caught, $1 - \sigma$ and the size of the penalty, θ but cannot avoid or insure against the risk of being caught. Let β_t be the fraction of income evaded in period t and let τ be the income tax rate at t. If the evader is discovered of evading an amount of income equal to $\beta_t W_2$, then the depositor has to pay a penalty on the unreported income in the same period t, but at a rate of θ, where $\theta > \tau$. So on receiving the endowment and in order to maximise his utility, the young-age depositor decides on: his consumption in both periods; β_t, the fraction of income to evade as well as d_t, the amount deposited at the bank (or his savings decision). After making his decisions, the *ex-post* tax state is revealed to the depositor. If the tax state is 'failure', the penalty is paid out of his savings.

Formally, the depositor must solve the following two-period problem:

$$\max_{c_{yt}, \beta_t, d_t, c^1_{ot+1}, c^2_{ot+1}} U = u(c_y) + \rho \sigma u(c^1_{ot+1}) + \rho(1 - \sigma)u(c^2_{ot+1}) \tag{5}$$

subject to:

$$p_t c_{yt} + p_t d_t \leq [\beta_t - (1 - \beta_t)(1 - \tau)]p_t W_2 \tag{6}$$

$$p_{t+1} c^1_{ot+1} \leq (1 + i_{dt+1})[d_t - \delta W_2]p_t + p_{t+1} W_3 \tag{7}$$

$$p_{t+1} c^2_{ot+1} \leq (1 + i_{dt+1})[d_t - \theta \beta W_2 - \delta W_2]p_t + p_{t+1} W_3 \tag{8}$$

$$0 \leq \beta_t \leq 1 \tag{9}$$

where $u(.) = log(.)$; $1 + i_{dt+1}$ is the gross nominal interest rate received in period t on deposits held by the banks; d_t are real deposits; c_{yt} is real young-age consumption; c^1_{ot+1} and c^2_{ot+1} is real old-age consumption in tax states

---

[6]The cost of evasion is not limited to only paying a penalty imposed by the government when the evader is caught, but it also includes cost of possible litigation, being excluded from certain public goods and even some social cost being regarded as a tax evader. For this model, however we will only consider a penalty as imposed by the government. The transaction cost that evading households incur, like hiring legal representatives or paying bribes to officials (Gupta and Ziramba, 2009) is accounted for through the depositor's old-age consumption function.

'success' or 'failure', respectively; ρ is the discount factor and δ represents the transaction cost that households incur to evade taxes. For clarity, (6) is the feasible first-period budget constraint, while (7) and (8) is the second-period budget constraint in the tax state where the depositor evades taxes successfully and where the depositor is discovered and incurs a penalty, respectively. The constraint in (9) is self-evident. In equilibrium, budget constraints (6) to (8) hold with equality since the depositor's utility function is increasing in consumption in each period. We define $1 + r_{dt+1} = \frac{1+i_{dt+1}}{1+\pi_{t+1}}$ as being the gross real interest rate on deposits held at banks. Solving the depositor's two-period utility maximisation problem yields the following first-order conditions (FOC):

$$d_t: \ u'(c_{yt}) = \rho(1 + r_{dt+1})[\sigma u'(c_{ot+1}^1) + (1 - \sigma)u'(c_{ot+1}^2)] \tag{10}$$

$$\beta_t: \ \tau_t u'(c_{yt}) \leq \rho\theta_t(1 - \sigma)[1 + r_{dt+1}]u'(c_{ot+1}^2) \tag{11}$$
$$\tau_t u'(c_{yt}) = \rho\theta_t(1 - \sigma)[1 + r_{dt+1}]u'(c_{ot+1}^2)$$
$$\tau_t u'(c_{yt}) \geq \rho\theta_t(1 - \sigma)[1 + r_{dt+1}]u'(c_{ot+1}^2)$$

for $\beta_t = 0$, $0 \leq \beta_t \leq 1$ and $\beta_t = 1$, respectively. From the series of first order conditions for β_t in (11), the left-hand side of the equation represents the marginal benefit of tax evasion and the right-hand side the marginal cost of tax evasion. The FOC's for the depositor imply that when the marginal cost of tax evasion exceeds the marginal benefit, there is no incentive for tax evasion so that $\beta_t = 0$. Conversely, when the marginal benefit of tax evasion exceeds the marginal cost, there is no incentive to declare any income so that $\beta_t = 1$. When the marginal benefit of tax evasion is equal to the marginal cost of tax evasion, there exist a range of plausible tax evasion parameters, such that $0 \leq \beta_t \leq 1$. However, for this interior solution to realise, it is required that $\tau_t > \theta_t(1 - \sigma)$ or that the regular tax rate is higher than the prospective penalty[7].

2.3 Financial intermediaries

There exist a finite number of risk-neutral banks in this economy, which we assume to behave competitively and are all subject to an obligatory cash reserve requirement, γ_t set by the government. This assumption assures that all banks levies the same cost on its loans, the gross nominal interest rate of $1 + i_{lt}$. In each period t, banks accept deposits and extend loans to risk-neutral entrepreneurs, subject to γ_t with the goal of maximising their profits. A simplifying assumption that deposits are one-period contracts assures a gross nominal deposit rate of $1 + i_{dt}$. Banks receive interest income from loans to entrepreneurs and meet their interest obligations to depositors at the end of the period. Because entrepreneurs have an incentive to

---

[7]Both Atolia (2007) and Sandmo (2012) provide a detailed account for this requirement.

declare bankruptcy even if their investment projects are successful, banks face a costly state verification problem, and hence offer a financing contract to entrepreneurs detailing the conditions of intermediation. Part of the conditions is that monitoring will take place if bankruptcy is declared. It is assumed that banks adopt a stochastic monitoring technology *à la* Bernanke and Getler (1989).

We denote λ as the number of times a misreporting entrepreneur can be discovered, with V the corresponding punishment. We use the revelation principle[8] to derive the optimal solution to the following financial contract based on the given structure. Formally, banks wish to maximise the following profit function:

$$\max_{i_l,L,V} \Pi_{Bt} = \frac{P_{t-1}}{P_t}[q(1+i_{lt})l_{t-1}+m_{t-1}-\lambda(1-q)cl_{t-1}-(1+i_{dt})d_{t-1}] \quad (12)$$

subject to:

$$l_{t-1} + m_{t-1} \leq d_{t-1} \quad (13)$$

$$m_{t-1} \geq \gamma_{t-1}d_{t-1} \quad (14)$$

$$q[p_t\alpha(W_1+l_{t-1})-p_{t-1}(1+i_{lt})l_{t-1}] \geq p_tq\alpha W_1 \quad (15)$$

$$q[p_t\alpha(W_1+l_{t-1})-p_{t-1}(1+i_{lt})l_{t-1}] \geq qp_t[\alpha(W_1+l_{t-1})-\lambda V] \quad (16)$$

$$p_tV \leq p_t\alpha(W_1+l_{t-1}) \quad (17)$$

$$0 \leq \lambda \leq 1 \quad (18)$$

where Π_{Bt} is the bank's profit at time point t; l_{t-1} is loans provided to entrepreneurs in period $t-1$; m_{t-1} is the bank's holding of fiat money; c is the bank's proportional cost for the monitoring technology and d_{t-1} is the deposits held by depositors at the bank in period $t-1$. The constraints (13) to (18) are explained as follows: (13) is the feasibility condition in order for the bank to satisfy its balance sheet; (14) is the legal reserve requirement obligating the bank's holding of fiat money; the 'participation constraint' ensuring that entrepreneurs accept the financing contract is given by (15); (16) is the 'incentive constraint' compelling entrepreneurs to not misreport the outcome of successful investment activities; (17) is the 'limited liability' constraint imposing a maximum penalty on entrepreneurs who misreport. Again, (18) is self-evident.

Solving the optimal contract for the financial intermediary requires (15) to be binding, leading to $\alpha = \frac{1+i_{lt}^*}{1+\pi_t}$. Incentive compatibility in (16) then requires $\lambda V = \alpha l_{t-1}$. Since the profit of the bank, Π_{Bt} decreases as monitoring increases, banks will set λ to its minimum such that (16) holds. Consequently, $0 < \lambda^* < 1$ and V is then set to its maximum, which from

---

[8]This induces entrepreneurs to truthfully report the outcome of their investment activity to the bank, as it is not more profitable to misreport the outcome, as reported in more detail in Myerson (1979).

(17) implies that $V^* = \alpha(W_1 + l_{t-1})$. This also ensures that (18) is binding. Then, assuming that entrepreneurs have no incentive to misreport because misreporting the actual outcome of the investment activity does not yield a higher expected profit to the entrepreneur, we ensure that (16) is binding and $\lambda^* = \frac{l^*_{t-1}}{W_1 + l^*_{t-1}}$. A competitive banking sector is characterised by free entry, which drives profits to zero. Thus, in equilibrium, based on the zero profit condition and that banks loan out all their available resources when $\alpha q > c$, we have that (13) and (14) also binds and hence, $l^*_{t-1} = (1-\gamma_{t-1})(d_{t-1})$. Besides from being an equilibrium condition, this also highlights the repressive nature of the obligatory reserve requirement in that it leads to sub-optimal functioning of the financial intermediary market.

So, given that $\alpha q > c$, the optimal financing contract is summarised as:

$$(i) \quad l^*_{t-1} = (1 - \gamma_{t-1})d_{t-1}$$

$$(ii) \quad \alpha^* = \frac{1 + i^*_{lt}}{1 + \pi_t}$$

$$(iii) \quad \lambda^* = \frac{l^*_{t-1}}{W_1 + l^*_{t-1}}$$

$$(iv) \quad V^* = \alpha^*(W_1 + l^*_{t-1})$$

2.4 Government

An infinity-lived consolidated government purchases g_t units of consumption goods, and government expenditure is assumed to be non-productive. The government finances its consumption expenditure through the collection of taxes, seigniorage income and penalty income that it levies on the unsuccessful depositor evading taxes. The government budget constraint is formally given by:

$$g_t = (1 - \beta_t)\tau_t W_2 + \frac{M_t - M_{t-1}}{p_t} + (1 - \sigma)\theta_t\beta_t W_2 \tag{19}$$

with the first part being the tax income, the second part being the seigniorage income (or inflation tax) in real terms and the third part being the penalty income it collects. Following Del Monte and Papagni (2001), we assume that the cost of monitoring tax evasion, say $(1 - \sigma)vW_2$, exactly offsets the penalty income derived from the evasion described in the third part of (19), so that the government budget constraint reduces to:

$$g_t = (1 - \beta_t)\tau_t W_2 + \frac{M_t - M_{t-1}}{p_t} \tag{20}$$

for simplicity. Also note that money evolves according to the following rule, $M_t = \mu_t M_{t-1}$ with μ_t the gross growth rate of money and $M_t = \gamma_t D_t$.

3   Equilibrium

A competitive equilibrium for this economy is defined as a sequence of prices $\{i_{lt}, i_{dt}, p_t\}_{t=0}^{\infty}$, allocations $\{c_{yt}, c_{ot+1}^1, c_{ot+1}^2, \beta_t, d_t\}_{t=0}^{\infty}$ as well as policy variables $\{\tau_t, \gamma_t, \theta_t, \mu_t, g_t\}_{t=0}^{\infty}$ such that:

- Given τ_t, θ_t, i_{dt} and $W_{i=1}^3$, the depositor optimally chooses β_t and savings, d_t;

- The equilibrium money market condition, $m_t = \gamma_t d_t$ holds for all $t \geq 0$;

- The loanable funds market equilibrium condition, $i_{lt} = \frac{i_{dt}}{(1-\gamma_t)}$ given the total supply of loans $l_t = (1-\gamma_t)d_t$, holds for all $t \geq 0$;

- Banks maximise profits subject to i_{lt}, i_{dt} and γ_t;

- The equilibrium resource constraint, $y_t = c_t + i_t + g_t$ holds for all $t \geq 0$, where $c_t = c_{yt} + qc_{ot+1}^1 + (1-q)c_{ot+1}^2 + C_{t+1}^e$ and $y_t = \sum_{i=1}^3 W_i$;

- The government budget constraint in (20) is balanced on a period-by-period basis;

- and d_t, m_t, i_{lt}, i_{dt} and p_t is positive for all periods.

4   Solving the model for the steady state degree of shadow economy

Taking the equilibrium conditions for this economic setting and imposing steady-state on the economy, thus no growth in the economy, we allow the government to follow time-invariant policy rules such that τ_t, γ_t, θ_t and μ_t are all constant over time and realising that in equilibrium $\pi = \mu$, or that the money growth rate equals the inflation rate, we yield a series of equations that allows us to solve the steady state model.

The depositor's optimisation solution essentially yields two equations: one for d^*, the steady state size of deposits in real terms and one for β^*, the steady state tax evasion parameter (or the steady state size of the shadow economy). Formally:

$$d^* = \frac{[(1+\rho)\tau - \theta(1+\rho(1-\sigma))][W_3 + \delta(1+r_d)W_2] + (1+r_d)W_2(\theta\rho\sigma(1-\tau))}{(1+r_d)(1+\rho)(\theta-\tau)} \tag{21}$$

and

$$\beta^* = \frac{\rho(\tau - \theta(1-\sigma))[W_3 + (1+r_d)W_2(1-\delta-\tau)]}{(1+r_d)W_2(1+\rho)(\theta-\tau)\tau} \tag{22}$$

From (5) it is also verified that $\frac{\partial^2 U}{\partial d} < 0$ and $\frac{\partial^2 U}{\partial \beta} < 0$, to ensure that both solutions are in fact, a maximum. It is evident from both (21) and (22) that the depositor's inter-temporal decision between making real deposits and evading taxes depends somewhat on $1 + r_d$, the gross real rate on deposits held at banks, besides from the real factors like θ, the penalty rate imposed by government when agents are caught evading taxes, τ, the tax rate imposed by government on the young-age endowment and σ, the probability of successfully evading taxes. Therefore, to understand the shadow economy behaviour in this setting it is crucial to understand exactly how $1 + r_d$ impacts the agent's tax evasion and savings decisions.

Firstly, we evaluate how both the real deposits and the fraction of income evaded change with observed changes in $1 + r_d$. For β^* we have:

$$(i) \quad \frac{\partial \beta^*}{\partial r_d} : -\frac{W_3 \rho(\tau - (1-\sigma)\theta))}{(1+r_d)^2(1+\rho)(\theta-\tau)} < 0$$

since $\tau > (1-\sigma)\theta$ was required to hold in order to obtain an interior solution for β^*, and for d^* we have:

$$(ii) \quad \frac{\partial d^*}{\partial r_d} : \frac{W_3[\theta(1+\rho(1-\sigma)) - \tau(1+\rho)]}{(1+r_d)^2(1+\rho)(\theta-\tau)} > 0$$

since $\theta > \tau$. Thus, in line with *a priori* expectation, β^* decreases with an increase in $1 + r_d$ and the size of real deposits, d^* increases with an increase in $1 + r_d$.

Secondly, from the bank's profit maximisation problem, we have:

$$1 + r_d = q\alpha(1-\gamma) + \frac{\gamma}{1+\mu} - \frac{\lambda c(1-q)(1-\gamma)}{1+\mu} \qquad (23)$$

where inflation has been set equal to the money growth rate, μ and λ is a function of $1 + r_d$ itself through the real deposits, d^*. From the optimal financing contract and substituting the loanable funds market equilibrium condition into the expression for λ^*, we have:

$$\lambda^* = \frac{(1-\gamma)d^*}{W_1 + (1-\gamma)d*} \qquad (24)$$

which together with both (21) and (23) yields an explicit expression for the gross real rate on deposits to analyse how financial development, which here is captured by both costly state verification c and λ, as well as inflation through μ, impact on the shadow economy in this model. Formally:

$$1 + r_d = q\alpha(1-\gamma) + \frac{\gamma}{1+\mu} - \left[\frac{(1-q)c(1-\gamma)}{1+\mu}\right]\left[\frac{(1-\gamma)d^*}{W_1 + (1-\gamma)d^*}\right] \quad (25)$$
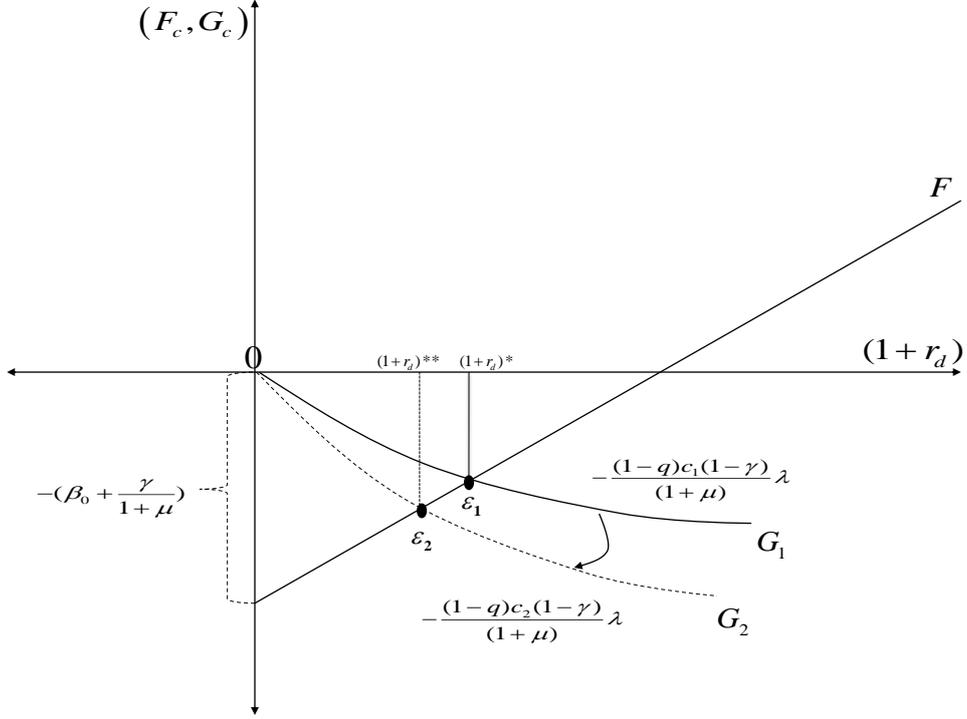
Figure 1: The effect of an increase in monitoring costs by the banks on the real deposit rate.

From (23) it should be evident that $1 + r_d$ is decreasing in c, the bank's monitoring cost parameter. However, to graphically illustrate the impact of the bank's monitoring cost on $1 + r_d$, we decompose $1 + r_d$ into two parts and then map the decomposed (23) in Figure 1.

The slope - and hence the shape - of the function is determined by the rate of change in λ. Near the zero bound interest rate, or where $1 + r_d$ approaches 0, the slope of the function approaches ∞ and as $1 + r_d$ approaches ∞, the slope of the function approaches 0. The function is therefore concave, increasing in $1 + r_d$, but at a decreasing rate.

Here $c_2 > c_1$, F is the intercept representation of the function and G_1, G_2 is the slope representation of $1 + r_d$ corresponding with the increase from $c_1 \to c_2$, respectively. As the banks' monitoring cost increases, there is a downward movement from G_1 to G_2 as the value of the slope increases. This increase in cost results in a new equilibrium level ε_2 which corresponds with the real deposit rate, $(1 + r_d)^{**}$ which is clearly lower than the initial equilibrium level ε_1 which corresponds with the real deposit rate, $(1 + r_d)^*$. The movement in the results presented here flow in the opposite direction for any given decrease in c.

The underlying intuition is straightforward: the higher the cost of mon-

itoring and the higher the incidence of the stochastic monitoring technology employed by the banks, the lower is $1 + r_d$. Conversely, higher CSV corresponds with a lower level of financial development, implying a lower level of incentives for the depositor to save and hence, a higher incentive for the depositor to evade in this setting. It is expected that $1 + r_d$ is decreasing in λ, as an increase in the probability (or number of times) that misreporting entrepreneurs can be discovered should lead to an increase in costs for the bank and therefore to a higher CSV altogether.

However, what is not immediately clear from (25) is the impact of inflation on $1 + r_d$. It should be clear, however that there are two competing effects of inflation on $1 + r_d$, one an *intercept* effect (from the first two terms in (23)) and the other a *slope* effect (from the last term in (23)). The amplified graphical representation in Figure 2 again separates these two effects to illustrate how the inflation or money growth rate parameter, μ, impacts both the intercept and the slope of the function $1 + r_d$ in (23) for the case of an increase in μ.
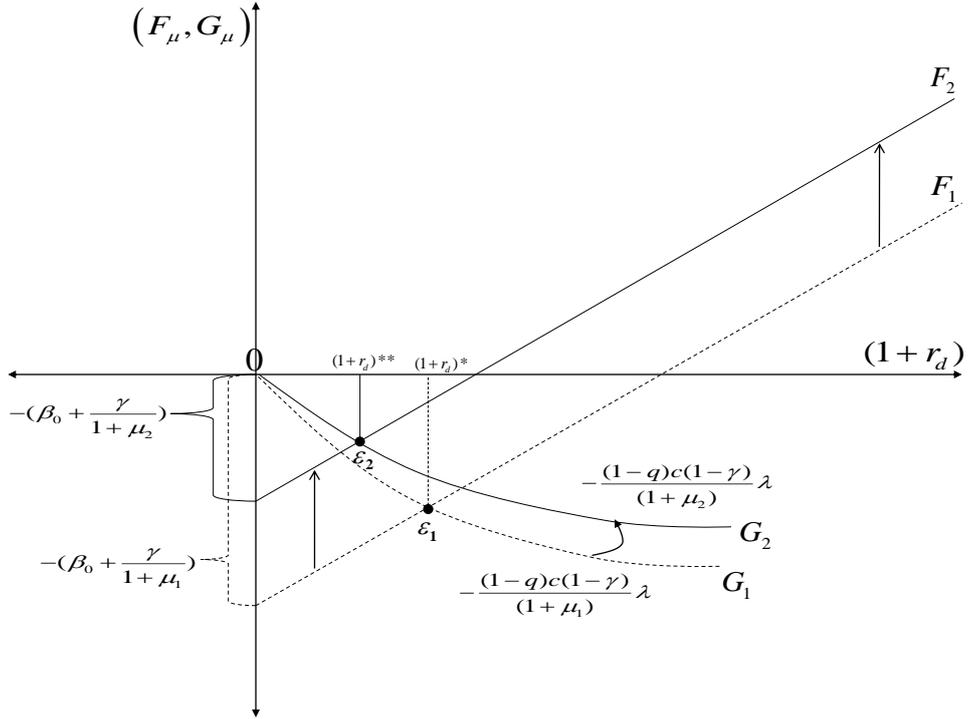

Figure 2: The dual effects of an increase in inflation on the real deposit rate.

Here $\mu_2 > \mu_1$, F_1, G_1 is the *intercept* and *slope* of $1 + r_d$ corresponding with μ_1, respectively and F_2, G_2 is the *intercept* and *slope* of $1 + r_d$ corresponding with μ_2, respectively. As inflation increases from μ_1 to μ_2, there is an upward shift in the intercept as the intercept becomes smaller,

13

from F_1 to F_2. Concurrently, there is an outward swing from G_1 to G_2 as inflation increases. This concurrent increase in inflation results in a new equilibrium level ε_2 which corresponds with the real deposit rate, $(1 + r_d)^{**}$ which is clearly lower than the initial equilibrium level ε_1 which corresponds with the real deposit rate, $(1 + r_d)^*$. These results hold for legitimate parametrisation as outlined in Gupta (2005) and Gupta and Ziramba (2008) for $\gamma > (1 - q)c(1 - \gamma)\lambda$. Taking the calibrated values for these parameters as found in these studies, we let: $c = 0.10$, $\gamma = 0.10$, $q = 0.64$, and lastly $\lambda = 0.10$, which yields $0.1 > (1 - 0.64) * 0.10 * (1 - 0.10) * (0.10)$ or $0.1 > 0.00324$. Thus, the required condition holds, and the outcome remains qualitatively equivalent for all plausible and legitimate values of the parameters. Chen (2003) provides further empirical support for the plausible values of specifically γ and c.

These results indicate that as inflation increases the real rate on deposits decreases, and from $\frac{\partial \beta^*}{\partial r_d}$ it would imply that β^* increases. So, as the depositor in this economy observes a decrease in the real rate on deposits held at banks, he decides to evade a bigger portion of his income leading to an increase in the size of β^*. In summary, (21) to (25) and the consequential analysis highlights the most important result that emerges from this analysis: that the fraction of income evaded by a depositor depends not only on real factors such as tax rates, τ; penalty rates, θ; and the probability of getting caught, $(1 - \sigma)$; but it also hinges critically on the monetary policy parameters in the model, namely the reserve requirement, γ and inflation, π as well as on the bank's cost parameters, c and λ.

5   The empirical setting

Recent empirical studies focus mainly on real factors as determinants of the size of the shadow economy. Fishlow and Friedman (1994) find that when current income decreases, tax compliance decreases and hence, the size of the shadow economy increases. Scneider (1994) shows how the imposed penalty rate leads to a higher shadow economy, while Schneider and Enste (2000) argue that the probability of being detected influences the size of the shadow economy. Dabla-Norris and Feltenstein (2005) show that the optimal tax rate may lead to a bigger shadow economy. Dreher, Kotsogiannis and McCorriston (2009) show that improved institutional quality decreases the size of the shadow economy and Elgin (2009) further argues that it is political turnover that determines the size of the shadow economy. Onnis and Tirelli (2011) argue that public expenditures decrease the size of the shadow economy, while Cerqueti and Coppier (2011) show how corruption affects the shadow economy. Most recently, Alm (2012) states that higher tax audit rates may reduce the size of the shadow economy and lastly Bose et al. (2012) provides evidence that it is the level of financial development

that determines the size of the shadow economy.

The focal point of a separate strand of the literature is the accuracy of different measures of the shadow economy. There are various different measures for the shadow economy - some more creative than others - but we will only highlight the most widely-used measures. Schneider, Buehn and Montenegro (2010) use a Multiple Indicators Multiple Causes (MIMIC) measure, which essentially is a structural equation model (SEM) with one latent variable. Thiessen (2010) constructs a shadow economy measure based on behavioural theories and Gomis-Porqueras, Peralta-Alva and Waller (2011) models the shadow economy using a currency demand or money demand approach. Onnis and Tirelli (2011) suggest using a Modified Total Electricity (MTE) approach and more recently, Elgin and Öztunali (2012) use a two-sector dynamic general equilibrium (DGE) model to obtain the size of the shadow economy. It should be mentioned that direct approaches, like surveys and structured questionnaires, are also widely used to obtain more "direct" measures of the size of the shadow economy.

The essence of our empirical testing however, is based exclusively on our theoretical framework.

5.1 Data

The data set used spans the period 1980 - 2009 and includes 150 countries[9], constituting a panel data set where $N = 150$ and $T = 30$. The period was chosen based on data availability for all key variables, and also to include at least one high and erratic inflationary period common in our panel - 1980 to 1990 - in the empirical analysis.

We compare two measures of the size of the shadow economy. $Shadow1$ is taken from the data set on the size of the shadow economy compiled by Schneider et al. (2010) using the MIMIC estimation method. This measurement covers the period 1999-2007. The second measure, $Shadow2$ is from a new data set compiled by Elgin and Öztunali (2012)[10], where the time-varying size of the shadow economy is estimated using a dynamic general equilibrium (DGE) model calibrated to a set of macroeconomic variables. This measure covers the period 1950-2009[11]. The correlation coefficient between $Shadow1$ and $Shadow2$, as stated in Elgin and Öztunali (2012) and also verified in this analysis, is 0.987. The strong correlation between these

---

[9]A list of all the countries included in this analysis is available from the authors.

[10]We gratefully acknowledge the use of the dataset on the shadow economy compiled by Ceyhun Elgin and Oğuz Öztunali.

[11]The econometric literature, in particular on the monetary model of exchange rate determination and purchasing power parity, suggests that it is the span of the data, and not the frequency that enhances econometric analysis of specifically long-run relationships between macroeconomic variables. This has been shown by Shiller and Perron (1985), Hakkio and Rush (1991), Otero and Simth (2000), Rapach and Wohar (2004) and more recently by de Bruyn, Gupta and Stander (2013).

two different measures of the shadow economy, based on different methodologies over different periods, facilitates balanced results as it excludes the way in which the measurements were calculated as a potential driver of the results.

$Bnkcost$ is a measure of the banking sector's average overhead cost, expressed as a percentage of the banking sector's total assets. Although this measure is not only restricted to the bank's monitoring cost parameter, as denoted by c in (23) of the theoretical model, by definition it includes c and is proposed here as a rational proxy for c in the absence of a more direct and widely available measure of c. This cost measure is also an indication of the efficiency with which commercial banks matches surplus units to deficit units in the economy, and is available from the Financial Structure dataset compiled and updated by Beck and Demirgüç-Kunt (2009). Barth, Caprio and Levine (2002) and Bose et al. (2012) have used $Bnkcost$ as a measure of the inefficiencies in the banking sector.

$Infl$ captures the effect of inflation and is the annual percentage change in consumer prices. As a proxy for inflation, we consider $Moneygr$ which is defined as the annual growth rate of the $M2$ monetary aggregate, since in steady state the money growth rate is set equal to the rate of inflation.

Cba, or central bank assets, is defined as the total claims that the central bank has on the domestic real non-financial sector and is expressed as a percentage of GDP. This variable is included as a relative measure of the size of the central bank in the economy and to account for the level of intervention - and the possible effect of financial repression - that economies experience. Curdia and Woodford (2011) extends a standard New Keynesian model and find central bank assets to be a factor in equilibrium determination. Bernanke and Reinhart (2004) also discuss the importance of central bank balance sheets and the composition thereof, in effectively implementing new and unconventional balance sheet policies. According to Christiano (2011), central bank intervention in asset markets may also prove to be very costly. This may lead to observing a higher overall banking cost in the economy. A 'bank balance sheet' channel for monetary policy through which the central bank can influence the loan decision of banks, was identified by Chami and Cosimano (2010). Through this channel the central bank may influence the bank's cost functions and ultimately, the decisions of the agent to deposit or evade.

The $Control$ variable set includes $Gdppc$, the real gross domestic product (GDP) per capita. Real GDP is a widely accepted measure of economic development in the literature (Boyd, Levine and Smith, 2001; Boyd and Jalal, 2012) and it may plausibly be used as an indicator of financial development since King and Levine (1993) showed that economic and financial development are closely related (Boyd and Jalal, 2012). The $Control$ variable set also includes two other important subsets, where the first set measures the level of financial development in each country, and comprises of: $Dcpb$,

the domestic credit provided by the banking sector; $Prvcrt$, the domestic credit provided by the banking sector as well as other financial institutions or intermediaries; $Intsprd$, the interest rate differential between loans and deposits; $M3$, the liquid liabilities as a percentage of real GDP and $Stmk$, a measure of stock market development calculated as the market capitalisation of all listed companies as a percentage of real GDP. These variables are often used in the financial development literature as indicators of the depth and the efficiency of both the banking and the financial sector (King and Levine, 1993; Levine and Zervos, 1998; Levine, Loayza and Beck, 2000; Boyd et al. 2001; Barth et al. 2002 and Boyd and Jalal, 2012). From these variables, we construct two financial development indicators using principal components analysis (PCA) and extract the unobserved common factors of these variables.

We define $Findev$ as the first proxy for financial development and it consists of the first principal component of the log-levels of $Dcpb$, $Prvcrt$, $M3$ and $Stmk$ which accounts for 80% of the variation in these four variables. We define the second proxy for financial development as $Findev2$, which consists of the first principal component of the log-levels of $Dcpb$, $Prvcrt$ and $Intsprd$ and it accounts for 68% of the variation in these three variables. $Intsprd$ is defined as the lending interest rate minus the deposit interest rate as published by the World Bank, and it indicates the magnitude of the wedge that financial repression induces between the interest rates that banks charge on loans and the interest rate banks offer on deposits[12]. This additional second proxy for financial development, given that $Gdppc$ is already a viable alternative to our first proxy, $Findev$, is an attempt to follow the recommendations of Levine (2005) and Boyd and Jalal (2012) that empirical measures of financial development should directly measure financial functions performed by the financial system. The PCA allows us to reduce the dimensionality of the set of variables to be included in our empirical analysis, whilst still retaining most of the informational content offered by these same variables (Bittencourt, 2012). It also aids in ensuring a more stable computational environment (Jolliffe, 1982).

The second subset measures the broad institutional quality of each country, and comprises of: $Regquality$, regulatory quality captures the perception of the ability of the government to formulate and implement sound policies and regulations that would permit and promote private sector development; $Ruleoflaw$ captures the perception of the extent to which agents have confidence in and abide by the rules of their respective society, in particular the quality of contract enforcement, property rights, the police and the courts, as well as the likelihood of crime and violence and $Fiscfreed$ measures fiscal freedom, or the extent of a country's total tax burden. All three

---

[12]Gupta (2005) provides a clear theoretical explanation of the characteristics of financial repression through obligatory high reserve requirements set by monetary authorities.

these variables are compiled as indices, with higher values of the index corresponding to better governance and a lower tax burden, respectively. These variables are commonly used in the shadow economy literature as important indicators of the policy, institutional and regulatory environment which impacts on the size of the shadow economy observed in countries (Schneider, 2007; Bose et al. 2012). Data on both *Regquality* and *Ruleoflaw* is from the World Governance Indicators (WGI) dataset maintained by the World Bank and covers the period 1996-2009, and *Fiscfreed* is found in the Index of Economic Freedom dataset compiled by The Heritage Foundation and covers the period 1995-2009. All other data is taken from the World Development Indicators and Global Development Finance (WDI) dataset published by the World Bank.

We again employ PCA to construct a proxy for the institutional, regulatory and policy strength of the countries in our sample. *Instit* is the first principal component of the *levels* of *Regquality*, *Ruleoflaw* and *Fiscfreed* and accounts for 66% of the total variation in these three variables. There is strong consensus in the literature on the shadow economy that institutions are very important in depressing the size of the shadow economy[13].

All main variables are expressed in logarithmic form. This is consistent with the depositor's life-time log-utility function, and as detailed in de Bruyn, Gupta and Stander (2013) this also allows more accurate analysis of the *relative* effect of the change in one variable on the change in another, which here is the relative effect of both financial development and inflation on the size of the shadow economy.

To illustrate the behaviour of the variables of interest in explaining the size of the shadow economy, we provide firstly the correlation matrix of the log-levels of the two main explanatory and other control variables on the size of the shadow economy in Table 1.

<<< Table 1 about here >>>

Firstly, there is a strong positive correlation between the two measures of the shadow economy, confirming the findings in Elgin and Öztunali (2012). Both our variables of interest, *Bnkcost* and *Moneygr*, are positively correlated with both measures of the shadow economy, *Shadow*1 and *Shadow*2 as expected. Financial development seems to have the expected attenuating effect on the size of the shadow economy, as does the level of and the relative change in the level of institutional quality. *Gdppc* is negatively correlated to the shadow economy, implying that societies that are more developed both financially and economically seems to be engaging less in underground economic activity. Lastly, *Cba* is positively correlated to the shadow economy,

---

[13]In different economic settings, Koreshkova (2006), Dreher, Kotsogiannis and McCorriston (2009), Elgin (2009) as well as Onnis and Tirelli (2011) all produce results supporting the attenuating effect of good institutions on the size of the shadow economy.

and negatively correlated to both measures of financial development, institutional quality and *Gdppc*. This suggests that the size of total claims the central bank has over the real domestic non-financial sector, or the stronger the ability of the central bank to intervene in the market, the more adverse conditions these markets face in general.

We also present the simple ordinary least squares (OLS) regression lines between our variables of interest and the size of the shadow economy in Figure 3.
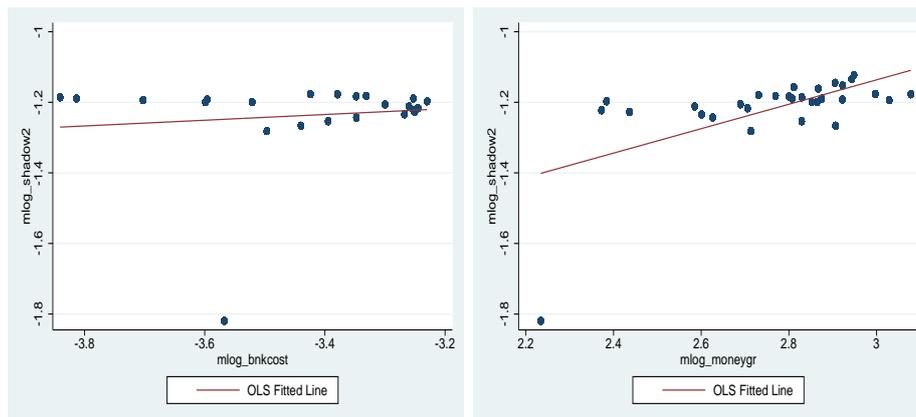

Figure 3: OLS regression lines of averages of the log-levels of Banking Cost and Money Growth on the Shadow Economy for 1980-2009, respectively.

Note the positive relationship between banking cost and money growth on the shadow economy[14]. From the observed data, it would seem that higher (lower) values of banking cost and higher (lower) values of the money growth rate, both correspond with higher (lower) values of the size of the shadow economy increases. Therefore - without implying causality at this stage - there seems to be some basis for our *a priori* expectation and predictions of our theoretical model, that both c and μ are positive with respect to the size of the shadow economy.

5.2   The empirical methodology employed

Since we have an unbalanced panel of observations from countries ($N = 150$) spanning multiple years ($T = 30$), and it is clear that there is some persistence in some of our variables of interest, we make use of dynamic panel (time-series) data analysis. The dynamic panel methodology allows us to deal more effectively with econometric problems like non-stationarity, joint statistical and economic endogeneity, potential simultaneity bias, unobserved country-specific effects that may lead to omitted variable bias and

---

[14]Both these positive relationships hold even when we use the short-span measure of the shadow economy, *Shadow*1.

importantly, measurement error. We are analysing the unobserved economy and as such measurement error is implied. This methodology also explores the added information from the time dimension in order to yield more accurate and informative estimates. Following Aghion, Bacchetta, Rancié and Rogoff (2009), we use the general method of moments (GMM) dynamic panel data estimator developed in Arellano and Bond (1991), Arellano and Bover(1995) and more specifically, the system GMM estimator developed in Blundell and Bond(1998). We compute Windmeijer-corrected two-step standard errors following the methodology proposed by Windmeijer (2005). This system GMM estimator addresses the aforementioned econometric issues in a dynamic formulation, where the lagged variable of the dependent variable is added to account for the persistence observed in the data[15].

We also expect our panel to be heterogeneous due to the inclusion of such a large number of countries with different economic, legal and regulatory policies, different political dispensations, different social issues which includes different levels of income inequality and different levels of both financial development and economic development. Moreover, the countries in our sample also share certain similar characteristics, like banking institutions, common monetary areas, trade agreements, monetary authorities and in some instances similar rules dictating their participation in the global economy. Our preferred estimator accounts for both scenarios.

Furthermore, within a dynamic panel environment, the stationarity of variables is always a concern. Not only is the stationarity assumption violated due to the persistence in the dependent variable series and - by design - the inclusion of its lagged value, but formal testing also indicates that the log-levels of $Bnkcost$ and $Gdppc$ are both $I(1)$. Using Fisher-type panel unit root tests that supports the well-known Phillips-Perron (1998) test for a unit root in the variable series, we compute a test based on the p-value of each one of the individual panels to test for stationarity. The value of the test statistic Z, an inverse Normal statistic, for $Bnkcost$ and $Gdppc$ is -0.481 and 10.00, respectively. Both these tests can not reject the null hypothesis of all panels containing a unit root, although further testing indicates that both variables are in fact, trend-stationary. In our case however, directly addressing the non-stationarity of variables is not required, since an additional advantage of the system GMM estimator is that because it uses differences of all the variables by design, it ensures that all the explanatory variables are stationary in any event (Baltagi, Demetriades and Law, 2009).

The dynamic benchmark empirical model that we will estimate, based on

---

[15]Roodman (2009) offers a step-by-step pedagogical account of the use of GMM style estimators.

our theoretical framework, is:

$$
\begin{aligned}
log\_shadow_{it} \;=\;& \beta_0 + \beta_1 moneygr_{it} + \beta_2 log\_bnkcost_{it} \qquad\qquad (26)\\
+\;& \beta_4 findev_{it} + \beta_5 log\_gdppc_{it} + \beta_6 instit_{it}\\
+\;& \beta_7 log\_shadow_{it-1} + u_{it}
\end{aligned}
$$

where log_shadow represents the measures of the size of the shadow economy, $moneygr$ is the money growth rate as a proxy for inflation, $log\_bnkcost$ is the proxy for the bank's monitoring cost, $findev$ represents the proxies for the level of financial development, log_gdppc is the real GDP per capita (also a proxy for the level of economic development in a country) and $instit$ represents the proxies for the institutional quality in this model. u is the error term that contains both country-specific as well as time-specific fixed effects, and of course a random disturbance term. Throughout the analysis, we will use both proxies for financial development.

5.2.1 Dynamic panel GMM estimation

Including the lagged dependent variable in (26), implies implicit correlation between the explanatory variables and the error term. This is because, by inclusion, the lagged shadow economy depends on u_{it-1} which contains the country-specific and time-specific effects. This further supports the choice of our preferred estimator suggested by Blundell and Bond (1998), which basically differences the model to get rid of country specific effects or any time-invariant country specific variable.

The moment conditions utilize the orthogonality conditions between the differenced errors and lagged values of the dependent variable. This assumes that the random disturbances term contained in u_{it}, are serially uncorrelated. We compute two diagnostic tests using the GMM procedure to test for first order and second order serial correlation in the disturbances. For validity, one should reject the null of the absence of first order serial correlation and not reject the null of the absence of second order serial correlation.

The dynamic system GMM estimation treats all the variables - other than the lagged dependent variable - as if they were either strictly exogenous or predetermined but not strictly exogenous, in that it assumes these variables are uncorrelated with the random disturbances in u_{it}. As *cautiously* stated in Baltagi et al. (2009), the differencing performed by the system GMM estimator may also remove any correlation due to the time-invariant common factors. An additional advantage of the system GMM estimator is that it does not "difference away" the fixed effects, but it instruments for the lagged dependent variable and other explanatory variables that may still be correlated with the disturbances by other variables believed to be uncorrelated with these fixed effects.

5.3 Empirical results

The two-step system GMM results reported in Table 2 provides encouraging support for the developed theoretical model. The recommendations of Roodman (2009) are followed and a detailed description of the different specification and instrument sets are provided first before the results are discussed.

$<<<$ Table 2 about here $>>>$

For columns (1) to (3), the endogenous (predetermined) or internal instrument set consists only of $log\_bnkcost$ and the lagged dependent variable. The external instrument set consists of log_cba, $log\_intsprd$ and the time-dummies for column (1); log_cba and the time-dummies for column (2) and only the time-dummies for column (3), respectively. This was done to reduce the instrument count and avoid proliferation of instruments. All instrument sets are valid, as supported by the Sargan statistic which is robust to heteroskedasticity and autocorrelation, where the hypothesis cannot be rejected that the instruments are exogenous. Moreover, the instrument count is always considerably less than the number of observations. The Arellano-Bond autocorrelation test also indicates that there is no serial correlation in the idiosyncratic disturbance term.

To account for the suspected economic endogeneity between $Moneygr$ and $Findev$ in a direct way, columns (1) to (3) exclude both principal components measures of financial development. In columns (4) to (6), both the different measures of financial development is included separately. Except for column (1), the forward orthogonal deviations transformation first suggested in Arellano and Bover (1995), are used as an alternative to the standard differencing. This transformation has the advantage of preserving sample size when the selected panel has gaps, or is unbalanced. All specification reported in Table 2 allow for the idiosyncratic disturbances to be both heteroskedastic as well as correlated within countries, but not across countries. Finite-sample Windmeijer (2005) corrected robust errors are reported in all columns.

The estimation performed in column (4) treats $log\_bnkcost$, $Findev$ and the lagged dependent variable of the shadow economy as endogenous, and therefore uses the second and deeper lagged values of these variables as internal instruments for the differenced equation and the first and deeper lagged values as instruments for the level equation. The size of the central bank, the interest rate spread and the time-dummies are treated as exogenous and therefore used as external instruments for the equation in levels. For column (5), $Findev$ is considered to be only predetermined and thus the first and earlier lagged values are used as instruments for the transformed equation and the difference of $Findev$ for the equation in levels. The external (or exogenous) instrument set remains the same. In column (6), the second

measure of financial development is again treated as an endogenous variable and included in the internal set together with $log\_bnkcost$ and $l.log\_shadow2$. The exogenous instrument set used in column (6) now excludes the interest rate spread, as this forms part of $Findev2$.

$Bnkcost$ estimates reported are positive and significant for all specified system GMM estimations. The size of the coefficients range from 0.00515 to 0.00872 and is in line with estimates obtained using the fixed effects (FE) estimator[16]. These results indicate that an increase in banking cost does indeed lead to an increase in the size of the shadow economy, suggesting that as banks face an increasingly costly state verification problem, the decision of the agents lean more towards evading a portion of their income and hence we observe an increase in the size of the shadow economy.

Moreover, with the exception of columns (3), (4) and (5), $Moneygr$ also presents positive and significant estimates against the size of the shadow economy. In the case of the exceptions, it is interesting to note that $Findev$ (and the second proxy, $Findev2$), or the measure of financial development, has an attenuating and significant effect on the size of the shadow economy. The results are consistent with those obtained by Blackburn et al. (2010) as well as Bose et al. (2012). The recent empirical work by Neely and Rapach (2011) on how common shocks, similar central bank functions and policy as well as international trade and capital flows produce commonality in international inflation, offers a more credible explanation for the not significant $Moneygr$ results reported in column (3), (4) and (5). The authors find that more than half of inflation variability in countries can be explained by international common influences that is not due to country-specific or time-specific effects in those countries, and since we are not controlling for the degree of openness observed in these countries we are not able to capture the inflationary effect in this specification. Bittencourt (2011) investigated the impact of inflation on financial development and concluded that low and stable inflation is a pre-condition for a more developed financial market. This supports the well-established linkage between financial development and inflation in the economic growth literature and more clearly documented in Boyd, Levine and Smith (2001), among others.

The estimates for $Gdppc$ are negative in those equations where a financial development measure is included, and positive in those equations where financial development was not explicitly modelled. Across all columns, $Gdppc$ is not significant. It was expected that in more developed societies agents would have less incentive to evade a portion of their income, which is not what is observed here. Bearing in mind that both $Bnkcost$ and $Gdppc$ is also used as a further indication of the level of financial development in the analysis presented here, economic endogeneity may be driving these unex-

---

[16]The GMM results accords well with results obtained from fixed effects (FE) estimators. The FE results are available from the authors upon request.

pected results. A more plausible explanation could be found in the results for the institutional framework. *Instit* estimates are negative and significant throughout the specification (with the exception of column (5)), but highly significant once the measures of financial development were not included. These results coalesce with the findings of Koreshkova (2006), Elgin (2009) and Onnis and Tirelli (2011), among others. In this empirical setting, it is clearly institutions and the level of financial development that impacts on the size of the shadow economy, and not the per capita income levels or the level of economic development.

The lagged dependent variable is positive and significant in all specification, as expected from the persistent nature of the shadow economy. The size of the lagged coefficient in most columns is high, again raising concerns about non-stationarity and hence, spurious regression results. In simulation studies, performed by Blundell, Bond and Windmeijer (2001), the efficiency and bias of system GMM estimates are compared to other estimators in the presence of highly persistent series, and found to improve upon both the precision and finite sample bias of other estimates. Moreover, Phillips and Moon (1999) formalised the idea that the cross-sectional information added in a panel framework provides more information, and therefore a clearer signal about the average long-run relation parameter, or the coefficient of the lagged dependant variable. The authors provide panel asymptotic theory which shows that the estimate for the coefficient on the lagged variable is consistent for persistent series, and hence spurious regression results in a non-stationary panel analysis is less problematic.

The reported results demonstrates that for different measures of the size of the shadow economy, controlling for the broad level - and the more "Levine-like" functional level - of financial development, taking the level of economic development into consideration through *Gdppc*, controlling for the level of institutional quality and the ability of the central bank to intervene in the economy, *Bnkcost* and *Moneygr* are important determinants for the size of the shadow economy. Moreover, an increase in any of these variables leads to a contemporaneous increase in the size of the shadow economy. Finally, the results are consistent with the theoretical propositions in Section 4: an increase in banking cost and an increase in the inflation rate, leads to an observed increase in the size of the shadow economy or to more underground economic activity.

6 Concluding remarks

We develop a theoretical model using an OLG framework consisting of depositors, entrepreneurs, banks and the government, to analyse the relationship between endogenously determined tax evasion as an indication of the size of the shadow economy, and both financial development and inflation.

Financial development is defined through the introduction of a CSV problem faced by banks as the lender to entrepreneurs in the economy. This CSV problem forces banks to employ monitoring technology and incur monitoring cost in order to observe the same outcome as the entrepreneur, which increases the banks' cost function and leads to a decrease in the real interest rate on deposits held by banks. Following the broad literature, societies with a higher (lower) level of financial development will have a lower (higher) cost of state verification. Entrepreneurs endogenously determine the portion of their income to misreport or under-declare to the bank, but face the price of doing so in the form of higher costs for access to and conditions of obtaining credit. These higher costs, or lower real rate on deposits and hence a lower level of financial development, provides an incentive to depositors to participate in tax-evasion activities as the marginal benefit of tax evasion is at least equal to the marginal cost thereof.

The empirical results provide consistent support for the theoretical findings. Once the level of both economic development and institutional quality is accounted for, concurrent with the size of the central bank and hence its ability to intervene in the economy, the reported estimates are evident of the fact that lower (higher) levels of financial development and higher (lower) inflation causes a bigger (smaller) shadow economy. Thus from a policy perspective, the role of financial development and lower rates of inflation in curbing the size of shadow economy is of paramount importance.

References

Aghion, P., Bacchetta, P., Rancière, R. and Rogoff, K.: 2009, Exchange rate volatility and productivity growth: The role of financial development, *Journal of Monetary Economics* **56**, 494–513.

Alm, J.: 2012, Measuring, explaining, and controlling tax evasion: lessons from theory, experiments, and field studies, *International Tax and Public Finance* **19**, 54–77.

Arana, O. M. V.: 2004, Economic growth and the household optimal income tax evasion. Discussion Paper No. 275, Department of Economics, Universidad Nacionale de Colombia.

Arellano, M. and Bond, S.: 1991, Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, *Review of Economic Studies* **58**, 277–297.

Arellano, M. and Bover, O.: 1995, Another look at the instrumental-variable estimation of error-components models, *Journal of Econometrics* **68**, 29–51.

Aruoba, S. B.: 2010, Informal sector, government policy and institutions, *Working Paper* . University of Maryland, Department of Economics, MD., USA.

Atolia, M.: 2003, An OLG model of tax evasion with public capital, *Working paper* . Florida State University, Tallahassee, FL., USA.

Baltagi, B. H., Demetriades, P. O. and Law, S. H.: 2009, Financial development and openness: Evidence from panel data, *Journal of Development Economics* **89**, 285–296.

Barth, J. R., Caprio, G. J. and Levine, R.: 2004, Bank regulation and supervision: what works best?, *Journal of Financial Intermediation* **13**(2), 205–248.

Beck, T. and Demirgüç-Kunt, A.: 2009, Financial institutions and markets across countries and over time: Data and analysis. World Bank Policy Research Working Paper No. 4943.

Bernanke, B. and Gertler, M.: 1989, Agency costs, net worth and business fluctuations, *American Economic Review* **79**(1), 14–31.

Bittencourt, M.: 2011, Inflation and financial development: Evidence from Brazil, *Economic Modelling* **28**(1-2), 91–99.

Bittencourt, M.: 2012, Financial development and economic growth in Latin America: Is Schumpeter right?, *Journal of Policy Modeling* **34**(3), 341–355.

Blundell, R. and Bond, S.: 1998, Initial conditions and moment conditions in dynamic panel data models, *Journal of Econometrics* **87**, 115–143.

Blundell, R., Bond, S. and Windmeijer, F.: 2001, *Nonstationary Panels, Panel Cointegration, and Dynamic Panels (Advances in Econometrics)*, Vol. 15, Emerald Group Publishing Limited, chapter Estimation in dynamic panel data models: Improving on the performance of the standard GMM estimator, pp. 53–91.

Bose, N., Capasso, S. and Wurm, M. A.: 2012, The impact of banking development on the size of shadow economies, *Journal of Economic Studies* **39**(6), 620–638.

Boyd, J. H. and Jalal, A. M.: 2012, A new measure of financial development: Theory leads measurement, *Journal of Development Economics* **99**, 341–357.

Boyd, J. H., Levine, R. and Smith, B. D.: 2001, The impact of inflation on financial sector performance, *Journal of Monetary Economics* **47**, 221–248.

Cerqueti, R. and Coppier, R.: 2011, Economic growth, corruption and tax evasion, *Economic Modelling* **28**, 489–500.

Chami, R. and Cosimano, T. F.: 2010, Monetary policy with a touch of Basel, *Journal of Economics and Business* **62**(3), 161–175.

Chen, B.-L.: 2003, Tax evasion in a model of endogenous growth, *Review of Economic Dynamics* **6**, 381–403.

Christiano, L.: 2011, Remarks on unconventional monetary policy, *International Journal of Central Banking* **7**(1), 121–130.

Cúrdia, V. and Woodford, M.: 2011, The central-bank balance sheet as an instrument of monetary policy, *Journal of Monetary Economics* **58**(1), 54–79.

Dabla-Norris, E. and Feltenstein, A.: 2005, The underground economy and its macroeconomic consequences, *The Journal of Policy Reform* **8**(2), 153–174.

de Bruyn, R., Gupta, R. and Stander, L.: 2013, Testing the monetary model for exchange rate determination in South Africa: Evidence from 101 years of data, *Contemporary Economics* **7**(1), 5–18.

DelMonte, A. and Papagni, E.: 2001, Public expenditure, corruption, and economic growth: the case of Italy, *European Journal of Political Economy* **17**, 1–16.

DiGiorgio, G.: 1999, Financial development and reserve requirements, *The Journal of Banking and Finance* **23**, 1031–1041.

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

Elgin, C. and Öztunali, O.: 2012, Shadow economies around the world: Model based estimates. Working Papers 2012/05, Bogazici University, Department of Economics, Turkey.

Fishlow, A. and Friedman, J.: 1994, Tax evasion, inflation and stabilisation, *Journal of Development Economics* **43**, 105–123.

Gomis-Porqueras, P., Peralta-Alva, A. and Waller, C.: 2011, Quantifying the shadow economy: Measurement with theory. Working Paper 2011-015A.

Gupta, R.: 2005, Costly state monitoring and reserve requirements, *Annals of Economics and Finance* **6**(2), 263–288.

Gupta, R.: 2008, Tax evasion and financial repression, *Journal of Economics and Business* **60**, 517–535.

Gupta, R. and Ziramba, E.: 2008, Costly tax enforcement and financial repression, *Economic Notes* **37**(2), 141–154.

Gupta, R. and Ziramba, E.: 2009, Tax evasion and financial repression: a reconsideration using endogenous growth models, *Journal of Economic Studies* **36**(6).

Gupta, R. and Ziramba, E.: 2010, Misalignment in the growth-maximizing policies under alternative assumptions of tax evasion, *The Journal of Applied Business Research* **26**(3).

Hakkio, C. S. and Rush, M.: 1991, Co-integration: how short is the long run?, *Journal of International Money and Finance* **10**(4), 571–581.

Holman, J. A. and Neanidis, K. C.: 2006, Financing government expenditure in an open economy, *Journal of Economic Dynamics and Control* **30**(8), 1315–1337.

Jolliffe, I. T.: 1982, A note on the use of principal components in regression, *Journal of the Royal Statistical Society* **31**(3), 300–303. Series C - Applied Statistics).

King, R. and Levine, R.: 1993, *Capital markets and financial intermediation*, Cambridge University Press, Cambridge, UK, chapter Financial intermediation and economic development.

Koreshkova, T. A.: 2006, A quantitative analysis of inflation as a tax on the underground economy, *Journal of Monetary Economics* **53**, 773–796.

Levine, R.: 2005, *Handbook of Economic Growth*, chapter Finance and Growth: Theory and Evidence, p. 865–934.

Levine, R., Loayza, N. and Beck, T.: 2000, Financial intermediation and growth: Causality and causes, *Journal of Monetary Economics* **46**, 31–77.

Levine, R. and Zervos, S. J.: 1998, Stock markets, banks, and economic growth, *The American Economic Review* **88**, 537–558.

Lucas, R. E.: 1976, Econometric policy evaluation: A critique, *The Phillips Curve and Labor Markets: Carnegie-Rochester Conference Series on Public Policy* **1**(1), 19–46.

Myerson, R.: 1979, Incentive-compatibility and the bargaining problem, *Econometrica* **47**, 61–73.

Neely, C. J. and Rapach, D. E.: 2011, International comovements in inflation rates and country characteristics, *Journal of International Money and Finance* **30**, 1471–1490.

Onnis, L. and Tirelli, P.: 2011, Institutions, policies and economic development. what are the causes of the shadow economy? Working paper series No. 206.

Otero, J. and Smith, J.: 2000, Testing for co-integration: power versus frequency of observation - further Monte-Carlo results, *Economics Letters* **67**(1), 5–9.

Phillips, P. C. and Moon, H. R.: 1999, Linear regression limit theory and nonstationary panel data, *Econometrica* **67**(5), 1057–1111.

Phillips, P. C. and Perron, P.: 1988, Testing for a unit root in time series regression, *Biometrika* **75**, 335–346.

Rapach, D. E. and Wohar, M. E.: 2004, Testing the monetary model of exchange rate determination: a closer look at panels, *Journal of International Money and Finance* **23**(6), 867–895.

Roodman, D.: 2009, How to do xtabond2: An introduction to difference and system GMM in Stata, *Stata Journal* **9**(1), 86–136.

Roubini, N. and Sala-i-Martin, X.: 1995, A growth model of inflation, tax evasion, and financial repression, *Journal of Monetary Economics* **35**, 275–301.

Sandmo, A.: 2012, An evasive topic: theorizing about the hidden economy, *International Tax and Public Finance* **19**, 5–24.

Schneider, F.: 1994, *Essays on Economic Psychology*, Springer Publishing Company, Berlin, Heidelberg, Germany, chapter Measuring the size and development of the shadow economy. Can the causes be found and the obstacles be overcome?, pp. 193–212.

Schneider, F.: 2007, Shadow economies and corruption all over the world: new estimates for 145 countries, *Economics* **9**.

Schneider, F., Buehn, A. and Montenegro, C. E.: 2010, New estimates for the shadow economies all over the world, *International Economic Journal* **24**(4), 443–461.

Schneider, F. and Enste, D.: 2000, Shadow economies: sizes, causes and consequences, *Journal of Economic Perspectives* **38**, 77–114.

Shiller, R. J. and Perron, P.: 1985, Testing the random walk hypothesis: Power versus frequency of observation, *Economics Letters* **18**(4), 381–386.

Sweidan, O. D.: 2011, Central bank losses: causes and consequences, *Asian-Pacific Economic Literature* **25**, 29–42.

Thießen, U.: 2010, The shadow economy in international comparison: Options for economic policy derived from an OECD panel analysis, *International Economic Journal* **24**(4), 481–5091.

Windmeijer, F.: 2005, A finite sample correction for the variance of linear two-step GMM estimator, *Journal of Econometrics* **126**(1), 25–51.

A   Appendix A

Table 1: Correlation Table

Variables	log_shadow1	log_shadow2	moneygr	log_bnkcost	findev	findev2	instit	log_instit	log_gdppc	log_cba
log_shadow1	1.000									
log_shadow2	0.993*	1.000								
moneygr	0.293*	0.279*	1.000							
log_bnkcost	0.493*	0.504*	0.266*	1.000						
findev	-0.680*	-0.666*	-0.427*	-0.585*	1.000					
findev2	-0.618*	-0.610*	-0.340*	-0.548*	0.918*	1.000				
instit	-0.727*	-0.717*	-0.331*	-0.351*	0.690*	0.673*	1.000			
log_instit	-0.628*	-0.557*	-0.213*	-0.147*	0.450*	0.448*	0.837*	1.000		
log_gdppc	-0.704*	-0.685*	-0.254*	-0.380*	0.679*	0.655*	0.802*	0.622*	1.000	
log_cba	0.285*	0.296*	0.100*	0.182*	-0.160*	-0.152*	-0.362*	-0.098*	-0.373*	1.000

1. A * indicates the 5% significance level.

Table 2: Estimation results of the System GMM Estimators

Variables	1	2	3	4	5	6
moneygr	0.00005**	0.00004*	0.00001	0.00001	0.00001	0.00004*
	(0.00002)	(0.00002)	(0.00002)	(0.00002)	(0.00002)	(0.00003)
log_bnkcost	0.00515*	0.00725**	0.0068*	0.00869**	0.00872**	0.00552*
	(0.00328)	(0.00378)	(0.00436)	(0.00356)	(0.00354)	(0.00299)
findev				-0.00729**	-0.00812***	
				(0.00280)	(0.00287)	
log_gdppc	0.00636	0.00767	-0.00230	-0.00268	-0.00224	-0.00270
	(0.01026)	(0.01095)	(0.02019)	(0.00681)	(0.00765)	(0.00783)
L.log_shadow2	0.93439***	0.91591***	0.899909***	0.92295***	0.91823***	0.91627***
	(0.02652)	(0.04071)	(0.03757)	(0.02495)	(0.02509)	(0.02631)
instit	-0.02740***	-0.02923***	-0.02955**	-0.01105*	-0.01170	-0.01744**
	(0.00932)	(0.01039)	(0.01310)	(0.00735)	(0.00805)	(0.00854)
findev2						-0.00468*
						(0.00302)
Obs	979	1,257	1,517	713	713	963
F	1,778.22	912.18	994.80	1,114.09	1,117.52	1,004.67
Sargan	0.62	0.58	0.61	1.00	1.00	1.00
AR1	0.001	0.001	0.001	0.001	0.001	0.002
AR2	0.442	0.439	0.531	0.638	0.605	0.310
#ofInstruments	143	142	141	165	164	172

1. A *, **, *** indicates the 10%, 5% and 1% significance level, respectively.

2. Robust standard errors in parenthesis.

3. All time dummies are suppressed to save space.