IMPACT OF VAT WITHHOLDING SCHEMES

Fernando Vásquez S. [a, b]

[a] Banco Central de Reserva del Perú
[b] Pontificia Universidad Católica del Perú

November 2017

Abstract

The schemes of advance payment through tax withholding constitute a widely used tool in Latin America to reduce tax evasion levels; however the studies on their effectiveness are very scarce and the few made are done in partial equilibrium analysis. Given the magnitude of the withheld amounts and the transversal nature of their application through different sectors in the economy, their impacts should be assessed in the context of a general equilibrium model. The present paper contributes to the literature by estimating a model of this type for the Peruvian economy, the methodology also allows to estimate internal evasion cost to the firm. The results show that the scheme constitutes an effective mechanism to reduce evasion and the implied distortions: a withholding rate at 2 percent can increase revenues by 13 percent.

JEL N° E7, H2,
Keywords: Tax Withholding, Tax Evasion, VAT, General Equilibrium

1. Introduction

One of the main problems of Latin American economies in fiscal matters is the low level of compliance with tax obligations, which translates into the low collection figures characteristic of the region. In response, during the last two decades, the main tax administrations[1] introduced withholding schemes as a mechanism to improve the productivity of value-added taxes.

The objective of these schemes is to ensure collection at those points in the production chain where there is greater informality. For this, a formal and "highly visible" taxpayer to the tax administration is designed to take charge of collecting the obligation of a third party, either of the previous stage (when acts as buyer) or of the later one (when acts as seller). As a result the tax administration could obtain resources from those "less visible" tax payers.

In Peru, withholding mechanisms were introduced in mid-2002 framed in three schemes: Retention Regime, Detracting Regime and Perceptions Regime. In the first two the agent appointed as a collaborator of the tax authority acts as a buyer, while in the latter act as a seller. In 2015, the amount collected under these schemes amounted to US$ 7 billion (3.8 percent of GDP), representing 45 percent of the annual revenue from the VAT.

However, the amounts collected through these schemes do not necessarily imply a net increase in revenue. The mechanism could merely be an advance payment[2], if the total tax obligations had been paid later on by the taxpayers. In order to generate a positive net effect, the tax withheld must be made on agents who otherwise would not have complied with their tax liability.

Despite the importance of these regimes, economic analysis about their functioning and effectiveness is very scarce. Moreover, the few studies made focus on partial equilibrium models, lacking the analysis of the impacts on the other macroeconomic variables or the negative effects on efficiency. The present paper analyzes the effectiveness of the advance payments mechanisms through the estimation of a dynamic model of general equilibrium.

The model considers a productive structure in two stages: a single final good producer who requires intermediate goods sold by a large number of small producers. The value added tax paid in the purchase of inputs, is later used as a tax credit by the final good producer at the time of calculating their tax liability. The total tax collection results from the amounts declared by the final producer as well as the sum of the declarations of the intermediate producers.

[1] One of the pioneers was the Federal Public Revenue Administration of Argentina. Currently these mechanisms are also applied in Colombia, Chile and Mexico.

[2] Except in situations of high inflation, where an advance would reduce the "Olivera-Tanzi" effect (the fall in tax revenues in real terms due to the presence of inflation and lags in the collection).

There are two types of intermediate good producers: good or bad taxpayers. The later underreport their tax obligation, taking advantage of their lack of visibility to the tax authority. The application of an advance payment scheme implies designating the producer of final goods as a withholding agent, in charge of collecting in advance the tax obligation of all input producers through a surtax on the VAT. Thus the net effect on the fiscal revenues will depend on the number of producers of inputs who were bad taxpayers.

The results show that the application of the advance payment scheme with a rate of 2 percent would increase tax revenues by around 19 percent, through a reduction in tax evasion, and also attenuating its efficiency losses as less resources are diverted from productive uses.

The rest of the paper is divided in four parts. Section 2 reviews the main theoretical approaches to tax evasion, emphasizing indirect taxation and withholding mechanisms. Section 3 describes the mechanisms of advance payment in the case of Peru. In section 4 is presented a theoretical model that analyzes the rationality of the evader and the impact of the withholding mechanisms on the collection. Section 5 shows the results of the estimation. Finally in section 6 the conclusions are presented.

2. Literature Review

2.1. Rationale of the Evader

The first analysis on tax evasion from a rational agent perspective were done by Allignam and Sandmo (1972), as an application of the more general research done by Becker (1968) regarding the economy of crime. Allignam and Sandmo (1972) present the rational choice of an adverse risk taxpayer under uncertainty, using a von Neumann-Morgenstern utility function. Thus the taxpayer decides if evade the payment of income tax taking into account a detection probability of undeclared income, and a penalty if detected.

A different perspective in the characterization of the rational evader is the one elaborated by Mayshar (1991) and more recently Slemrod (2001), which introduce a tax concealment technology. Because concealment implies an effort or the employment of resources that negatively affect the generation of profits, evasion will be determined by an internal optimum, in which the cost of hiding income with tax savings are equalized.

2.2. Evasion of Indirect Taxes

The theoretical analysis of evasion has focused mainly on the personal income tax, the number of investigations at the level of companies is smaller, and even less are those directed to the analysis of indirect taxes.

A first analysis of indirect tax evasion was done by Marrelli (1984), in the context of a monopolist that faces an advalorem tax rate on sales. The model allows to analyze, in addition to the evasion itself, the degree of tax traslation to the consumers. The results show that evasion does not affect the traslation of the tax.

Virmani (1989) uses a competitive equilibrium approach in his analysis of evasion of indirect taxes. The model assumes: i) that the probability of detection increases with production; Ii) that the function of tax concealment depends on the proportion of the declared sales. As a result of evasion, productive inefficiency occurs (evasion affects production) and it depends positively on tax rates.

2.3. Evasion and taxes withheld

In order to reduce the costs of collection, the tax administration designates agents to withhold certain taxes that correspond to tax obligations of a third party (note that this equivalent to an advance payment of the obligation).

Yaniv (1988) presents the first analysis of evasion in the context of a tax witholding agent. The model considers the case of a competitive entrepreneur who is responsible for the retention of a percentage of the salary of their workers, deciding whether or not to correctly state the amount of wages paid by the company. However, the amount evaded by this mechanism can raise the amount to be paid by income tax (profits).

In the case of witholding between firms the literature is quite limited. The first study that analyzes indirect taxation withholding was made by Keen (2008). The model raises the case of VAT withholdings as an instrument to obtain resources and reduce the tax evasion of an informal sector that does not declare its sales. Accordingly, the economy has two sectors: formal and informal, which compete in the production of a final consumption good, using an imported input. The withholding mechanism is established as a VAT surtax on input imports, and their use is analyzed in comparison to the application of tariffs.

The use of tariffs allows tax collection from the informal sector, even if they do not declare its sales they must at least pay taxes when importing their inputs. Keen's analysis shows the usage of a tariff with this purpose is a inferior solution than a VAT with a withholding mechanism. In the case of the surtax to the VAT rate (retention rate) also allows to obtain resources from the informal sector, but due to the creditable nature of the VAT, the formal sector can deduct the tax paid. This implies an advantage of the withholding mechanism above a tariff, as the latter is not creditable, it implies an increase in the costs not only in the informal sector but also in the formal one.

Later on, Carrillo, Emran and Rivadeneira (2012) and Brockmeyer and Hernández (2016) analyze the effects of withholdings in the case of domestic sales (not imports). In both cases, they propose a partial equilibrium model for a rational evader in the tradition of Allignam and Sandmo (1972), where the application of a retention mechanism leads to taxpayers bunching around a threshold determined by the rate of withholding (they declare taxes just to the limit of what is withheld). To corroborate empirically these

predictions, they use microeconomic information from individual taxpayer tax files, from Ecuador in the case of Carrillo et al (2012), and from Costa Rica in Brockmeyer and Hernández (2016).

In the case of Brockmeyer and Hernández (2016) they also report an increase in the tax declared amounts, as a result of the application of the withholding mechanism. However, the Costa Rican scheme is designed as an advance payment of the income tax instead of VAT, and the greater revenues are obtained by the withholdings on self-employed workers than at the level of firms.

The present paper extends the literature on the effectiveness of an advance payment scheme in the context of a general equilibrium model, trying to identify the effects of the scheme on aggregate collection and macroeconomic variables. According to the literature review, no general equilibrium analysis has been performed so far. Unlike the previous studies, the behavior of the evader is characterized by the Mayshar (1991) and Slemrod (2001) concealment approach, that is more appropriate than the Allignam and Sandmo tradition, in the case for risk-neutral agents, as would be firms. The methodology also allows to estimate evasion cost internal to the firm.

3. The advance payment scheme in Peru

Since mid-2002, as part of a package of tax measures to increase tax collection, three administrative mechanisms for advance payment of tax obligations are applied: Retentions, Detractions, and Perceptions regimes.

The objective of these mechanisms is to ensure that the tax authorities capture revenue generated in sectors where there are high levels of evasion. These sectors are characterized by a significant atomization of taxpayers, which reduces the effectiveness of individual control actions such as those carried out on large taxpayers. Therefore, large economic units, which carry out frequent transactions with small taxpayers, were selected to withhold a certain amount as an advance payment on account of the taxes that correspond to them.

3.1. **Retention**

Under this mechanism, some buyers are designated as "Retention Agent" by SUNAT[3] and are required to retain 3% of the sale price (as a VAT surtax) of goods, first sale of real estate, services and construction contracts taxed with VAT, whose invoice amount exceeds S /. 700. The amount withheld must be transferred to the tax authorities along with the rest of the own tax obligations of the "Retention Agent".

[3] Superintendencia Nacional de Aduanas y de Administración Tributaria (SUNAT) is the tax administration institution in Peru.

Accordingly, suppliers (sellers of goods, service providers or builders) receive a "proof of withholding" that indicates the amount withheld, which can be applied against the payment of their obligation for VAT, or if there is no obligation, request its devolution.

The usual criterion for designating retention agents is to be a large buyer who make transactions with many small suppliers. Both the designation and the exclusion will be made through a Resolution of the SUNAT. To date, 4248 taxpayers have been designated as "retention agents".

3.2. **Perception**

Under this regime, some sellers (usually of supplies or intermediate goods) are designated as "Perception Agents", which must apply a surtax over the sale price (perception). Currently this mechanism also applies to certain imports, in which case, the agent of perception is SUNAT itself. The amount received will be transferred by the seller to the tax authorities, and it will be deducted by the purchaser his VAT obligation to pay. Perception does not apply if the buyer is an end consumer. However, a withholding applies to people who purchases goods for an amount equal to or above S /. 1 500 (US$ 450).

To date, perception applies to three groups of transactions : Imports, Fuel Sales and domestic sales of certain goods.

3.3. **Detraction**

Through this mechanism, the buyer of a good or service included in the scheme withholds a percentage of the selling price and deposits it in an account at the Banco de la Nación (Estate Bank) under the name of the seller. Subsequently, the seller may use these funds to pay the tax debts (including fines and interest), contributions to ESSALUD and ONP; As well as costs and administrative expenses generated by other SUNAT actions to coercitive collection.

To date, there are four schemes of Detraction Regime: Domestic sales of goods and services (13 goods and 10 services), land transportation service, public ground transportation service and the Sale of Pilfered Rice (IVAP). A further description of the application of this regimen is shown in the Appendix.

3.4. **Tax Withholding Collections**

The resources collected by these systems show a growing trend, reaching a level equivalent to 4.6 percent of GDP (equivalent to 53 percent of VAT collection). Regime of detractions that obtains greater resources. During 2015, a drop of 0.8 percent of GDP has

been observed which could be associated to the reduction in withholding rates and the exclusions of some goods and services of the mechanism.

Table 1
Tax Withheld Amounts
(Millions of S/)

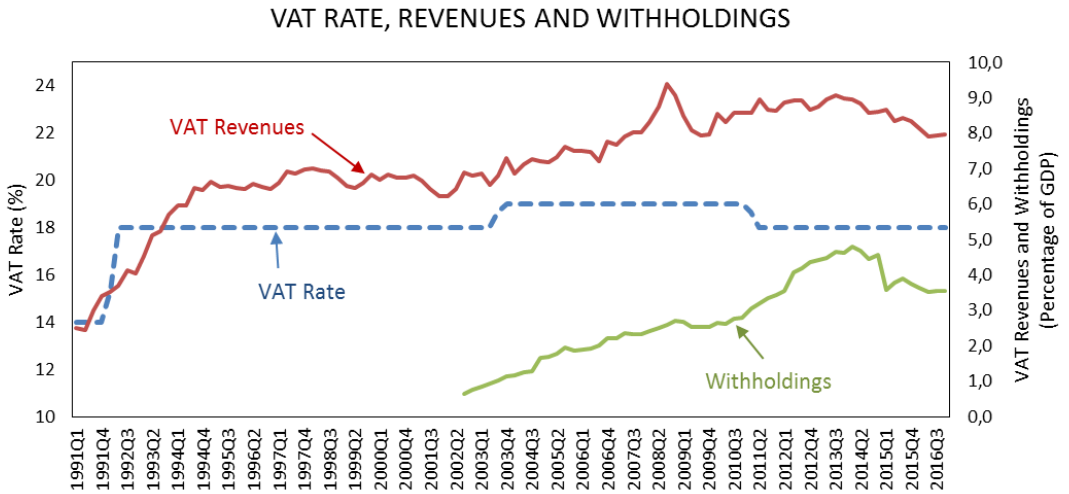
	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Retenciones	622	1 424	1 422	1 372	1 569	1 709	2 151	2 195	2 612	2 411	2 255	1 652	1 073	879
Detracciones	68	498	1 228	2 503	3 479	4 765	5 799	6 045	7 375	11 278	16 314	20 953	23 427	20 537
Percepciones	8	105	432	634	756	901	1 155	1 108	1 337	1 532	1 656	1 757	2 122	1 645
Total	**698**	**2 027**	**3 082**	**4 509**	**5 803**	**7 375**	**9 105**	**9 347**	**11 325**	**15 222**	**20 224**	**24 363**	**26 623**	**23 061**
% PBI	*0,4%*	*1,0%*	*1,4%*	*1,8%*	*2,0%*	*2,3%*	*2,6%*	*2,6%*	*2,7%*	*3,2%*	*4,0%*	*4,5%*	*4,6%*	*3,8%*
Part. % IGV	*6%*	*14%*	*19%*	*25%*	*27%*	*29%*	*29%*	*32%*	*32%*	*38%*	*46%*	*51%*	*53%*	*45%*
Rec. IGV	**12 612**	**14 116**	**16 203**	**18 302**	**21 517**	**25 258**	**31 587**	**29 520**	**35 536**	**40 424**	**44 042**	**47 819**	**50 352**	**51 668**
% PBI	*6,6%*	*6,9%*	*7,1%*	*7,4%*	*7,5%*	*7,9%*	*8,9%*	*8,1%*	*8,5%*	*8,6%*	*8,7%*	*8,8%*	*8,8%*	*8,5%*

Source: SUNAT, BCRP

In the case of Retention and Perception regimes, the amounts withheld have a direct effect on the collection of VAT, while in the case of Detractions, the amounts are not transferred immediately to the Public Treasury, since they are deposited in an account on behalf of the detracted agent.

However, the amounts collected do not necessarily imply an increase in the tax collection since part of it would have been paid later by the taxpayers.

Graph 1



4. Some (stylized) facts and policy effectiveness test

Simple plots of VAT revenues and Withholding series, suggest a close relation between both variables (see Graph1). It is possible to carry out an initial evaluation of the impact of prepayment measures on income through the counterfactual analysis developed by

Pesaran and Smith (2016). This methodology evaluates the conditional estimates of a reduced form that assumes that the policy is not implemented:

$$y_t = \Psi_x(\rho)x_t + \Psi_x(\rho)w_t + v_t$$

Where y is the target variable, x is the vector of policy variables, w is the vector of other control variables that are invariant to policy changes. A policy change from ρ^0 to ρ^1 will cause the equation to exhibit a break in T_0:

$$y_t = \Psi_x(\rho^0)x_t + \Psi_x(\rho^0)w_t + v_t \qquad t = M, M+1, M+2, \dots, T_0$$

$$y_t = \Psi_x(\rho^1)x_t + \Psi_x(\rho^1)w_t + v_t \qquad t = T_0+1, T_0+2, \dots, T_0+H$$

Under the null hypothesis that: i) no policy changes are made, ii) there are no changes in the other parameters, the counterfactual prediction of the target variable and, h periods thereafter will be:

$$\hat{y}_{T_0+h} = \Psi_x(\rho^0)\hat{x}_{T_0+h} + \Psi_x(\rho^0)w_{T_0+h} \qquad h = 1,2, \dots, H$$

The policy effects will be given by:

$$d_{T_0+h} = y_{T_0+h} - \hat{y}_{T_0+h}$$

Pesaran and Smith (2016) show that it is possible to test the null hypothesis given that:

$$X^2_{d,H} = \frac{\hat{d}'\hat{d}}{\hat{\sigma}^2_{0v}} \sim^a \chi^2_H$$

Where $\hat{d}$ is an array of the d_h. alternatively:

$$X^2_{d,H} = \sqrt{H}\frac{H^{-1}\sum_H d_h}{\hat{\sigma}^2_{0v}} \sim^a N(0,1)$$

Estimation results from pre and post policy intervention are shown in Table 2. The objective variable is VAT revenues as a percentage of GDP, policy variables are the VAT tax rate, the withholdings as a percentage of GDP. Control variables are the output gap and growth rate of the GDP deflator.

Policy effect show that from one monetary unit of tax withheld, up to 0.6 monetary units goes to a VAT revenue increase.

Table 2

REDUCED FORM EQUATION OF POLICY OBJECTIVE (VAT REVENUES) 1/

	Pre Intervention IQ 1991-IVQ2002	Post Intervention IQ 2003-IV 2016
Policy Variables		
VAT rate	0,354 **	0,585 **
	(0,080)	(0,169)
Withholdings		0,630 **
		(0,076)
Control Variables		
Inflation (GDP Deflator)	-0,106 **	-0,059
	(0,013)	(0,038)
Output Gap	12,935 **	12,065 **
	(1,899)	(4,279)
R-Squared	0,904	0,811

** Significant at 1%, * Significant at 5%.

1/ Cointegration estimation using Fully Modified OLS. Objective variable is VAT revenues as a percentage of GDP.

Results for the Pesaran-Smith (2016) tests are shown in Table 3, clearly the null hypothesis of no effectiveness is rejected.

Table 3

POLICY EFFECTIVENESS TESTS

	Value	Probability
$X_H^2 =$	776,57	0,00
$t_{d,H}^a =$	24,75	0,00

H= 56 observations

5. The Model

The structure of the economy is similar to that used by the new Keynesian macroeconomic models[4] using a two-stage production scheme: a large number of intermediate goods producers and final good sector. This basic structure is modified to introduce both tax evasion and the application of an advance payment mechanism on

[4] See for example Chari, Kehoe and McGrattan (2000) or Christiano, Eichenbaum and Evans (2005)

purchases from input suppliers. In order to maintain greater similarity with the design of the withholding mechanism, it is considered that small producers of intermediate goods are not visible to the tax authority, while a large producer of final goods acts as a withholding agent.

5.1. Consumers

The economy has a representative consumer facing the following utility maximization problem:

$$Max\ U = \sum_{t=0}^{\infty} \beta^t \left(\frac{C_t^{1-\sigma}}{1-\sigma} - \frac{L_t^{1+\eta}}{1+\eta} \right), \qquad (1)$$

subject to:

$$(1+\tau_t)P_t C_t + B_t = W_t L_t + (1+r_{t-1})B_{t-1} + Div_t\ , \qquad (2)$$

Where C_t is the consumption good (in charge of the firm producing final goods), L_t is the labor effort that is assigned to producers of intermediate and final goods, W_t the nominal wage, B_t are bonds (of a period of duration), Div_t are the dividends received from companies producing intermediate goods, and τ_t the VAT rate[5]. The parameter β measure the subjective discount rate, while σ and η determine the relative risk aversion of the utility función.

5.2. Production

There are two types of products, final goods and inputs. The production of final goods is carried out by a firm operating in perfect competition, while the input producers consist of a large number of small firms in monopolistic competition, which sell the intermediate good to the only producer of final goods. Unlike the canonical New Keynesian model, the production of final goods requires not only inputs but also labor, this implies that the producer generates value added that will be subsequently taxed by a VAT. This is important for the empirical counterpart, because most of the revenues in Peru come from large formal enterprises.

The final goods producer uses the following production function:

$$Y_t = Min\left\{ [\int Y_t(i)^q di]^{\frac{1}{q}}, \frac{L_t^Y}{f_t} \right\}, \qquad 0 < q \leq 1 \qquad (3)$$

Where $Y_t(i)$ is the i-th input, while L_t^Y constitutes the labor used in the production of the final good (f_t determines the unitary labor requirement), while *1/(1-q)* is the elasticity of substitution between inputs. The firm maximizes profits:

[5] Note that the general formulation of the problem considers the possibility of modifying the tax rates, so that they could have a different dynamic from the retention rates.

$$(1+\tau_t)P_tY_t - \int(1+\tau_t)P_t(i)Y_t(i)di - W_tL_t^Y - \tau_t[P_tY_t - \int P_t(i)Y_t(i)di], \quad (4)$$

Where the last bracket shows the VAT payments, which can be written as

$$P_tY_t - \int P_t(i)Y_t(i)di - W_tL_t^Y\,, \qquad (5)$$

So the demand for inputs will be given by:

$$Y_t^d(i) = \left[\frac{P_t - f_tW_t}{P_t(i)}\right]^{\frac{1}{1-q}} Y_t\,, \qquad (6)$$

While the demand for labor will be given by:

$$L_t^Y = f_tY_t\,, \qquad (7)$$

the superscript Y identifies the firm as a producer of the final good. As the firm operates in perfect competition the producer will not generate profits (zero profits):

$$P_tY_t = \int P_t(i)\left[\frac{P_t - f_tW_t}{P_t(i)}\right]^{\frac{1}{1-q}} Y_t di + W_tf_tY_t\,, \qquad (8)$$

so that:

$$P_t = \left[\int P_t(i)^{\frac{q}{q-1}}di\right]^{\frac{q-1}{q}} + f_tW_t, \qquad (9)$$

On the other hand the producers of intermediate goods are distributed in the interval [0,1], the i-th firm has the following production function:

$$Y_t(i) = \frac{L_t^I(i)}{s_t}, \qquad (10)$$

Where s_t indicates the labor force requirements per unit of product, and the superscript I identifies the firm as a producer of inputs. Producers seek to maximize their profits (Note that the whole value of their production is value added):

$$\pi_t(i) = (1+\tau_t)P_t(i)Y_t(i) - s_tW_tY_t(i) - \tau_tP_t(i)Y_t(i)\,, \qquad (11)$$

which can be expressed as:

$$\pi_t(i) = [P_t(i) - s_tW_t]\left[\frac{P_t}{P_t(i)}\right]^{\frac{1}{1-q}} Y_t\,. \qquad (12)$$

Taking the first-order condition with respect to $P_t(i)$ gives:

$$\left[\frac{P_t}{P_t(i)}\right]^{\frac{1}{1-q}} Y_t - \frac{1}{1-q}[P_t(i) - s_tW_t]\left[\frac{P_t}{P_t(i)}\right]^{\frac{q}{1-q}} Y_t \frac{P_t}{(P_t(i))^2} = 0\,, \qquad (13)$$

Which can be simplified to:

$$P_t(i) = \frac{1}{q} s_t W_t \qquad (14)$$

The price of the intermediate good is fixed considering a mark-up of *1/q* with respect to the nominal cost. In a symmetric equilibrium (all firms have the same production function and face a similar demand), so $P_t(i)=P_t(j)$, and:

$$P_t = (\frac{1}{q} s_t + f_t) W_t, \qquad (15)$$

5.3. Government

The government maintains a budget balance, financing spending with VAT and lump sum taxes (transfers). In a context of no evasion, the budget constraint will be given by:

$$(1+\tau_t)P_t G_t = \tau_t[P_t Y_t - \int P_t(i)Y_t(i)di] + \tau_t \int P_t(i)Y_t(i)di + Transf_t = \tau P_t Y_t + Transf_t, \qquad (16)$$

G_t are the units of the final good demanded by the government. The first equality on the right side reflects the sum of what is collected by the producer of the final good after deducting the tax credit and what is collected by the producers of inputs. The second equality reflects the characteristic that the sum of payments made by both producers equals what consumers pay for the final good.

5.4. Macroeconomic Equilibrium

The conditions for macroeconomic equilibrium will be given by:

$$\begin{gathered} U_t^C = \beta(1+r_t)U_{t+1}^C, \\ \frac{U_t^L}{U_t^C} = \frac{W_t}{(1+\tau_t)P_t}, \\ L_t^{Oferta} = L_t^Y + \int L_t^I(i)di = (f_t + s_t)Y_t, \\ C_t + G_t = Y_t, \\ P_t = (\frac{1}{q} s_t + f_t) W_t, \\ (1+\tau_t)P_t G_t = \tau_t P_t Y_t + Transf_t, \end{gathered} \qquad (17)$$

5.5. Introducing Tax Evasion

Small firms can avoid paying VAT through a "concealment technology." Nevertheless, this technology implies a monetary cost. The objective of the i-th intermediate goods producer will be:

$$\pi_t(i) = (1+\tau_t)[P_t(i)Y_t(i)] - W_t L_t^Y(i) - \tau_t[P_t(i)Y_t(i) - O_t(i)] - R_t(i), \qquad (18)$$

Where $O_t(i)$ corresponds to the undeclared (hidden) tax amounts, and $R_t(i)$ to the costs incurred to evade taxes given the concealment technology. It is assumed that both are functions of the labor L^O used to evade[6].

$$O_t(i) = O\left(L_t^O(i)\right) \qquad where\ O' > 0\ y\ O'' < 0\,, \tag{19}$$

$$R_t(i) = P_t(1+\tau)(\phi_i + \phi_0) + \delta W_t L_t^O(i) \qquad where\ \delta > 0\,, \tag{20}$$

Where ϕ_i is a idiosyncratic fixed cost to the i-th producer (in terms of the final good), distributed in the interval [0,1]. The evasion generates inefficiencies of two types, the producer must make expenses in terms of the final good and also he must use labor force to hide his income (or to justify expenses). Thus equation (18) can be rewritten as:

$$\pi_t(i) = [P_t(i)Y_t(i) - s_t W_t Y_t(i)] + \tau_t O\left(L_t^O(i)\right) - P_t(1+\tau)(\phi_i + \phi_0) - \delta W_t L_t^O(i)\,, \tag{21}$$

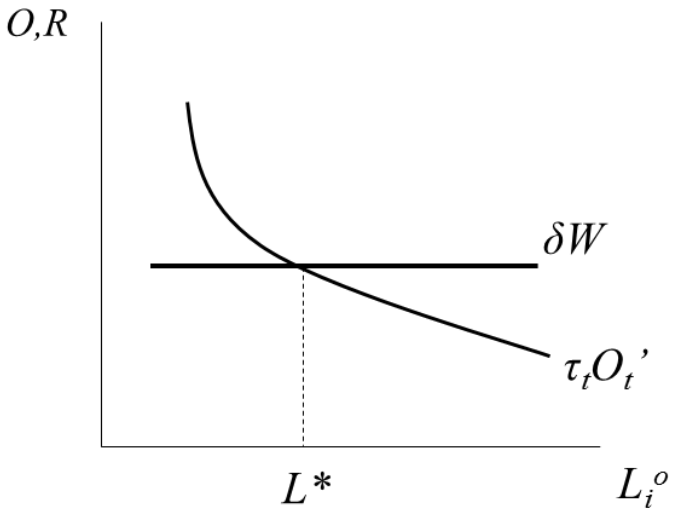
Therefore the firm will evade taxes until:

$$\tau_t O_t' = \delta W \tag{22}$$

$$\text{as long as:} \qquad \tau_t O\left(L_t^O(i)\right) > P_t(1+\tau)(\phi_i + \phi_0) + \delta W L_t^O(i)$$

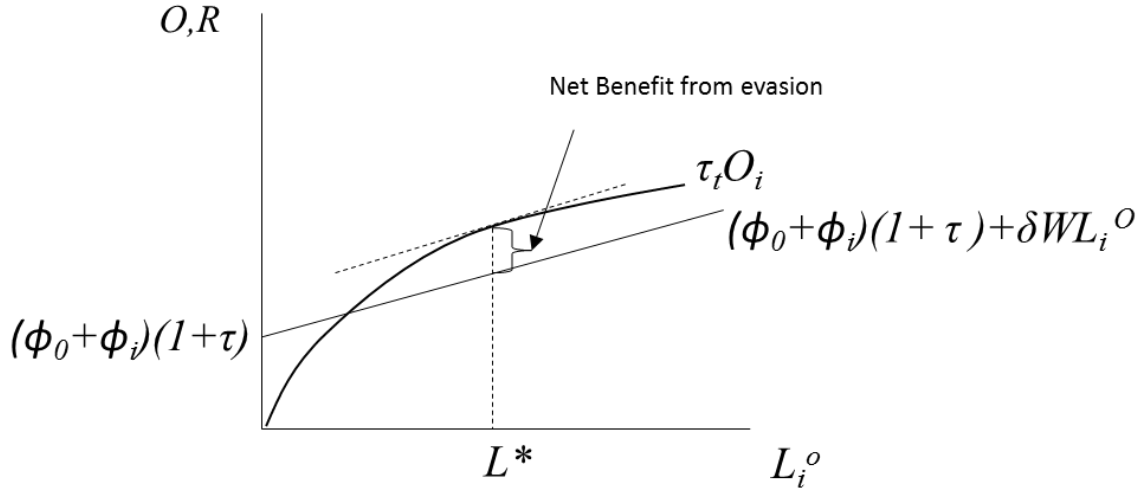
Thus, the producer of intermediate goods will under-report his tax obligation until the marginal benefit of evading ($\tau_t O_t'$) is equal to the marginal cost (δW). Figure 1(a) shows the optimal decision (22), graph 1(b) shows this result in levels.

Figure 1(a)



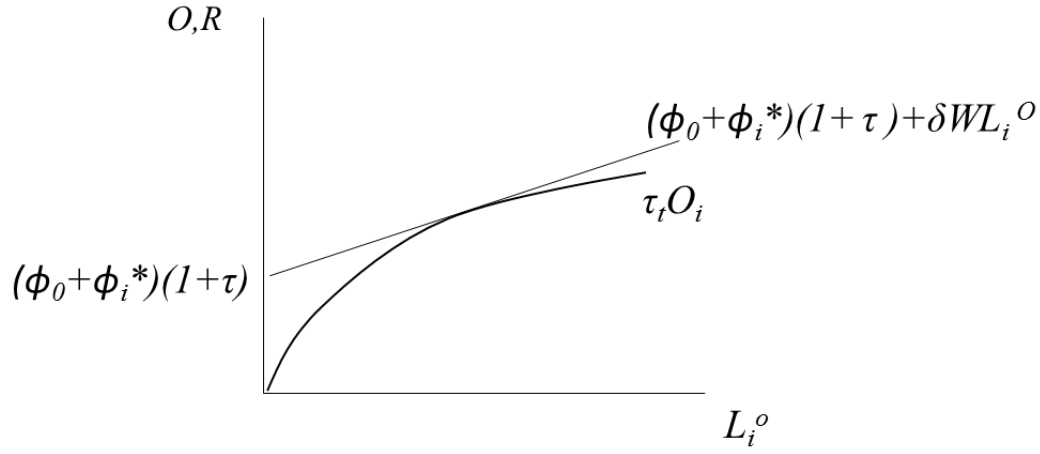
[6] Note that $O(i)$ has an upper bound: $O(i)<P(i)Y(i)$ as tax payments cannot be negative.

Figure 1(b)



If the firms are ordered according to its concealment cost, it is feasible to determine the *i-th* firm that it is indifferent to evade, figure 2 shows the solution for the marginal evader (firms with values of ϕ_i above ϕ_{i^*} do not evade).

Figure 2



The government's budget constraint is modified to:

$$(1+\tau_t)P_tG_t = \tau_t\left[P_tY_t - \int P_t(i)Y_t(i)di\right] + \tau_t\int_0^{i^*}[P_t(i)Y_t(i) - O_t(i)]di + \tau_t\int_{i^*}^{1}P_t(i)Y_t(i)di \quad (23)$$

Which could be written as:

$$(1+\tau_t)P_tG_t = \tau_tP_tY_t - \tau_t\int_0^{i^*}O\left(L_t^O(i)\right)di, \quad (23')$$

Note that the firm producing final goods does not alter its tax behavior. Input-producing firms continue to sell their product including the amount of the tax on the price, even though only a fraction of them ($i>i^*$) comply with it completely.

In turn, the equilibrium in the labor market will now be:

$$L_t^{Supply} = L_t^Y + \int L_t^Y(i)di + \int_0^{i^*} L_t^O(i)di\,, \qquad (24)$$

The last term measures the use of workers to hide rents.

5.6. Applying a withholding mechanism

In order to ensure collection (reduce evasion), the government decides to designate the final goods producer (good taxpayer) as a withholding agent for input producers.

If the withholding rate is equal to k_t, the profits of the i-th producer of intermediate goods will be:

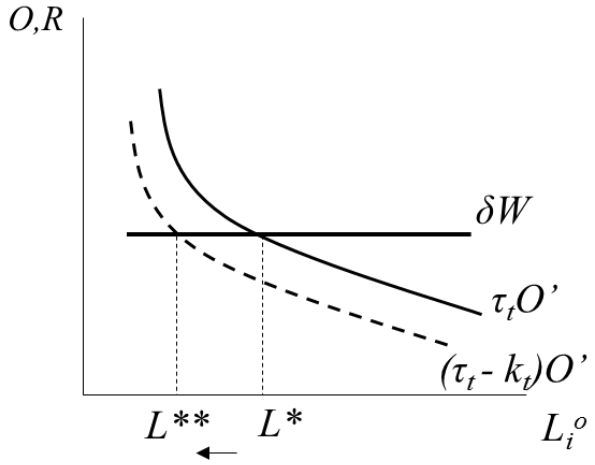
$$\pi_t(i) = (1+\tau_t - k_t)[P_t(i)Y_t(i)] - W_t L_t^Y(i) - (\tau_t - k_t)[P_t(i)Y_t(i) - O_t(i)] - R_t(i)\,, \quad (25)$$

Which is simplified to:

$$\pi_t(i) = P_t(i)Y_t(i) - W_t L_t^Y(i) - (\tau_t - k_t)O_t(i) - R_t(i)\,, \qquad (26)$$

implying a reduction in the amount evaded, graphically:

Figure 3



Likewise, the level of specific costs that makes the evader indifferent is reduced, so the number of companies that will sub-declare their tax liability is reduced. This can be seen in figure 3, by reducing the benefit of evading, L^* would move down to a level as L^{**}.

In this case the government's budget constraint will be:

$$(1+\tau_t)P_tG_t = \tau_t[P_tY_t - \int P_t(i)Y_t(i)di] + k_t \int P_t(i)Y_t(i)di + (\tau_t - k_t)\int_0^{i^{**}}[P_t(i)Y_t(i) - O_t(i)]di + (\tau_t - k_t)\int_{i^{**}}^1 P_t(i)Y_t(i)di + Transf_t \qquad (27)$$

Which can also be written as:

$$(1+\tau_t)P_tG_t = \tau_tP_tY_t - (\tau_t - k_t)\int_0^{i^{**}} O_t(i)di + Transf_t \ , \qquad (27')$$

While the new labor market balance:

$$L_t^{Oferta} = L_t^Y + \int L_t^Y(i)di + \int_0^{i^{**}} L_t^O(i)di \ , \qquad (28)$$

If tax concealment technology is defined as:

$$O_t(i) = P_tL_t^{O^v}(i) \qquad v \in [0,1] \qquad , \qquad (29)$$

The evader will use labor resources to under report tax by:

$$L_t^O(i) = \left[\frac{v}{\delta}(\tau_t - k_t)(\frac{1}{q}s_t + f_t)\right]^{\frac{1}{1-v}}, \qquad (30)$$

Clearly:

$$\frac{\partial L_t^O(i)}{\partial \tau_t} > 0 \quad , \frac{\partial L_t^O(i)}{\partial k_t} < 0, \qquad (31)$$

That is, an increase in the tax rate will raise tax evasion, while a higher rate of retention will decrease it. This expression also reflects the negative association between the output gap and the empirically observed evasion[7] (s_t and f_t are the labor requirements per unit of output so that the inverse of a positive shock to productivity).

Using equation (22), it is also possible to characterize the marginal evader as the firm with:

$$\phi_{i^{**}} = \frac{1}{(1+\tau_t)}(\frac{1}{q}s_t + f_t)^{\frac{v}{1-v}}\left\{(\tau_t - k_t)^{\frac{1}{1-v}}\left[\left(\frac{v}{\delta}\right)^{\frac{v}{1-v}} - \delta\left(\frac{v}{\delta}\right)^{\frac{1}{1-v}}\right]\right\} - \phi_0 \ , \qquad (32)$$

Again,

[7] Banco Central de Reserva del Perú (2017)

$$\frac{\partial \phi_{i^{**}}}{\partial \tau_t} > 0 \quad , \qquad \frac{\partial \phi_{i^{**}}}{\partial k_t} < 0,$$

This result allows the following observations:

i. A withholding scheme can reduce inefficiencies in the economy by eliminating the costs that producers of intermediate goods use to hide their tax liability. As less resources will be used to evade, the demand for labor will be reduced, the final effect on the labor market will depend on the income effect on labor supply resulting from higher taxes.

ii. It is also possible that the prepayment measure does not generate collection effects. This will occur when the costs of concealment exceed the benefit of evading. In the model this will occur when ϕ_i exceeds ϕ^* for any producer of intermediate goods.

6. Estimation methodology and Results

6.1. Estimation

The model is estimated with Bayesian methods described by Smets and Wouters (2003), using Metropolis-Hastings algorithm based on Monte Carlo simulations of Markov chains. For this purpose, the Dynare 4.4.3 program was applied to the data of the Peruvian economy in the period 2004-2016, with a quarterly frequency.

The estimation includes the definition of four exogenous variables (shocks): the VAT rate, the withholding rate, the unit labor requirements of producers of intermediate goods and the final good.

Four observable variables are considered, GDP, VAT revenues (domestic revenues net of mining and hydrocarbons) and withholdings. All series were turned stationary using the Hodrick Prescott filter.

The information of the priors and posteriors, the confidence intervals and the distributions used are reported in Table 4, while the graphs of the distributions are included in Appendix 2.

The results show that, in general, the parameters are well identified, with evasion technology (v) and costs associated with it (δ, ϕ_0) being especially relevant. A notable fact is that the estimate shows well the persistence of shocks in the rate of retentions, in which the downward revisions made by the economic authority have been registered after long periods of application of the scheme.

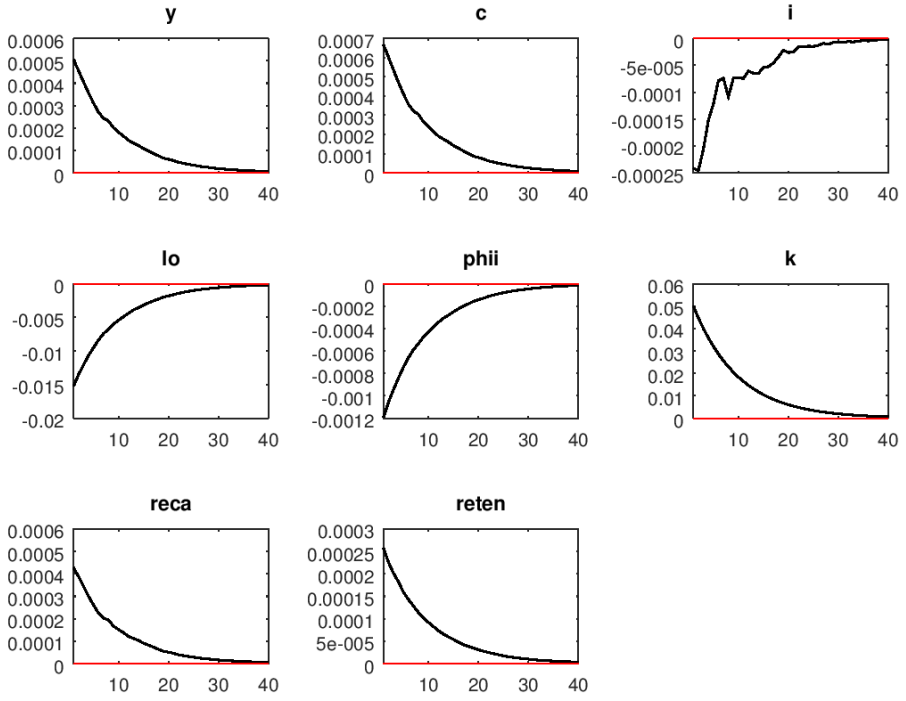
Table 4
PARAMETERS

	Prior mean	Post. mean	90% HPD Interval		Prior	pstdev
			Inferior	Superior	Distribution	
v	0.20	0.27	0.25	0.29	beta	0.03
ϕ_0	0.30	0.07	0.04	0.10	norm	0.05
δ	0.10	0.06	0.06	0.07	beta	0.01
β	0.95	0.96	0.94	0.97	beta	0.02
σ	0.90	0.95	0.89	0.97	beta	0.03
q	0.90	0.93	0.91	0.95	beta	0.03
η	1.00	1.01	0.92	1.03	beta	0.03
ρ_s	0.60	0.62	0.61	0.66	beta	0.02
ρ_f	0.50	0.51	0.46	0.55	beta	0.02
ρ_τ	0.20	0.24	0.16	0.29	beta	0.02
ρ_k	0.85	0.89	0.89	0.90	beta	0.02
ρ_g	0.80	0.80	0.77	0.84	beta	0.02

6.2. Simulation

Figure 4 shows the impacts of an innovation on the withholding rate

Figure 4

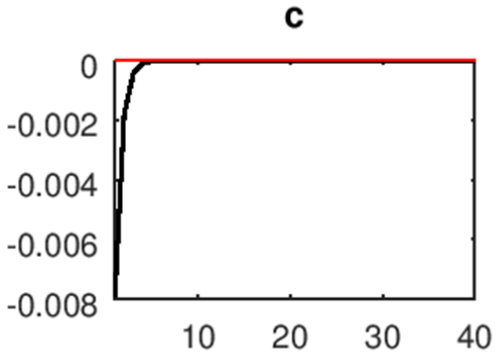


The results show, not only an increase in revenue, but also a reduction of the labor force used to evade taxes and a reduction in the proportion of companies producing intermediate goods that evade (ϕ_{ii}*). There is also an increase in household consumption, which is due to the fact that lower evasion reduces inefficiencies and costs for the economy.

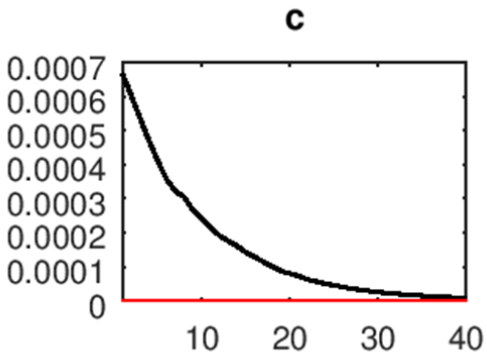
Indeed, the impacts on consumption are different in the face of an increase in the rate of VAT and an increase in the rate of withholdings, given an increase of one percentage point in the rate of each of the taxes (an increase in the withholding rate reduces de diversion of resources from productive uses while an increase in VAT rate reduces consumption possibilities):

Figure 5

Rise in VAT Rate



Rise in Withholding Rate



Finally, the model allows quantifying the effect on collection of introducing the retention scheme. Thus, considering an average retention rate of 2 percent, collection increases not less than 13 percent at steady state.

Table 5
Long Run Effects (Steady State)

	Withholding Rate		
	0%	2%	3%
y	0.497	0.507	0.512
c	0.326	0.340	0.347
r	0.043	0.043	0.043
g	0.160	0.160	0.160
L^o	2.008	1.708	1.564
ϕ_{ii}	0.092	0.068	0.057
Revenues	0.069	0.079	0.082
Withholdings	0.000	0.005	0.007

7. Conclusions

The present paper contributes to the literature by estimating the effects of the retentions scheme within the framework of a dynamic model of general equilibrium with tax evasion.

The introduction of the withholding scheme reduces the incentives to evade, being an effective tool to raise the collection from low visibility agents.

Also, since tax evasion generates inefficiencies in the economy, its reduction can improve the welfare of the population.

The results show that the introduction of a withholding scheme with a rate of two percent allows to increase tax revenues by about 13 percent.

APENDIX 1

To date, there are four Schemes of Withholding Regime: Internal sale of goods and services (13 goods and 10 services), land transportation service, passenger public ground transportation service and the Sale of Pilfered Rice (IVAP)

- Domestic sale of goods and services: buyers who purchase goods or contract services in amounts above S /. 700, detract from the sale price a percentage of draw. To date, 13 goods and 10 services are allocated to the withdrawal.

Goods included in the regime:
- Rate 15%: Industrial waste and scrap.
- Rate 10%: Sand and stone, gold taxed with IGV, non-auriferous metallic minerals, and non-metallic minerals.
- Rate 4%: Hydrobiological resources, hard yellow corn, fishmeal, wood, meats and edible offal, and real property taxed with IGV.
- Rate 1,5%: Goods exonerated from IGV, gold and other metals exempted from the IGV.

Services:
- Rate 10%: Labor intermediation, leasing of goods, maintenance and repair of movable property, movement of cargo, other business services, commercial commission, manufacture of goods by order, transportation of people, and other services taxed with IGV.
- Rate 4%: Construction contracts.

- Land transportation service: the user of the service when the price exceeds S /. 400 will deduct 4% of the amount of the transaction, which must be deposited in the account of the Bank of the Nation in the name of the service provider.

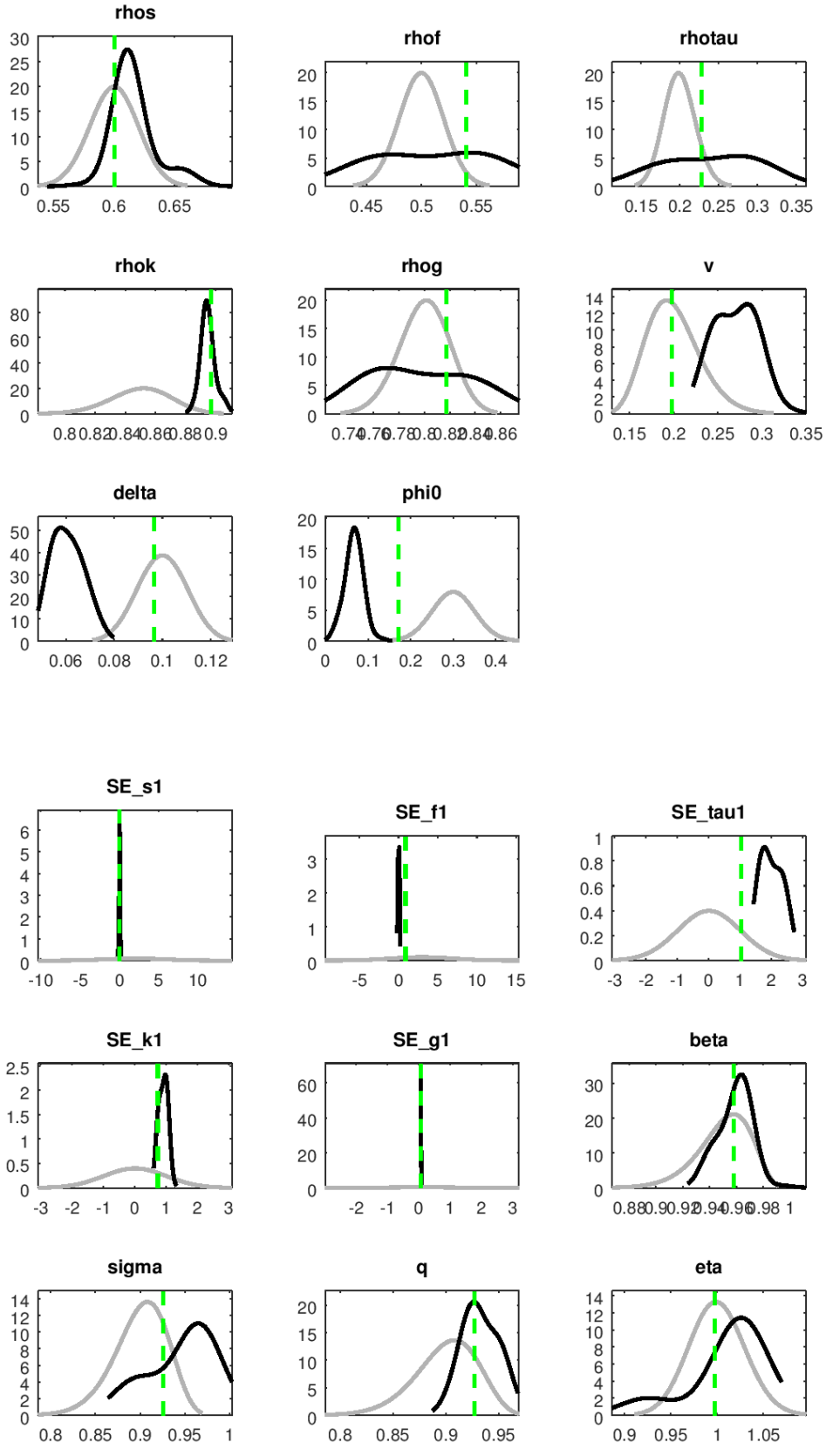
- Public passenger ground transportation service: the carrier is obliged to pay the amount of the toll fee to the Toll Administrator authorized by SUNAT, at the time the toll payment is to be made. The amount of drawdown is:

    - S /. 2 for each axis of the vehicle, in the case of toll booths or toll points that charge toll rates in both directions of transit.
    - S /. 4 for each axis of the vehicle, in the case of checkpoints or toll points that charge toll rates in one direction only.

The amount deducted will be deposited in the account of the Bank of the Nation in the name of the carrier.

For operations subject to the Sales Tax of Pilfered Rice (IVAP): in the first sale of pounded rice, the buyer must deduct 3.85% of the amount of the operation, provided that said amount exceeded S /. 700. The amount withheld will be deposited in a Bank of the Nation account in the name of the seller.

APENDIX 2


rhos
rhof
rhotau
rhok
rhog
v
delta
phi0
SE_s1
SE_f1
SE_tau1
SE_k1
SE_g1
beta
sigma
q
eta

References

Allingham, M.G., and A. Sandmo (1972), "Income tax evasion: a theoretical analysis", Journal of Public Economics 1(3-4), 323-338.

Arias, R.J. (2005), "A Note on Indirect Tax Evasion". Mimeo.

Banco Central de Reserva del Perú (2016), Reporte de Inflación. Setiembre.

Barbone, Bird, Vazquez (2012), "The Costs of VAT: A review of the Literature". Case Paper 106/2012.

Becker, G.S. (1968), "Crime and punishment: an economic approach", Journal of Political Economy 76(2),169-217.

Brockmeyer, A., Hernandez, M. (2016), "Taxation, Information, and Withholding". Policy Research Working Paper 7600, World Bank.

Carrillo, P., Emran, M., Rivadeneira, A. (2012), "Do Cheaters Bunch Together? Profit Taxes, Withholding Rates and Tax Evasion". Mimeo.

Chari, Kehoe, Mc Grattan (2000), "Sticky Price Models of the Business Cycles: Can the Contract Multiplier solve the persistence problem". Econometrica, Vol 68. September.

Christiano, Eichenbaum, Evans (2005), "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy". Journal of Political Economy, Vol. 113, 1.

Keen, M. (2008). "VAT, tariffs, and withholding: Border taxes and informality in developing countries," Journal of Public Economics, vol. 92(10-11), 1892-1906.

Keen, M., Smith, S. (2007) "VAT Fraud and Evasion; What Do We Know, and What Can be Done?" IMF Working Paper WP/07/31.

Lema, Rubén (2002). Estimación de la Evasión del IGV. Mimeo Banco Central de Reserva del Perú.

Marchese, C. (2006) "Taxation, Black Markets and other unintended Consequences". Chapter 10 in Handbook of Public Finance. Springer.

Marrelli, M. (1984). "On Indirect Tax Evasion", Journal of Public Economics 25, 181-196.

Mayshar, J. (1991), "Taxation with costly administration", Scandinavian Journal of Economics 93(1):75-88.

Pesaran, H., Smith, R. (2016) "Counterfactual analysis in macroeconometrics: An empirical investigation into the effects of quantitative easing". Research in Economics, Volume 70, Issue 2, June, 262–280.

Sivasakkaravarthi, P., Ganesan, D. (2011). "A Study on Literature Review on Value Added Tax". International Journal of Business Management, Economics and Information Technology. Vol. 3, No. 2, July-December 2011, 309-316.

Smets, F y Wouters, R. (2003). "An estimated dynamic stochastic general equilibrium model for the euro area". Journal of the European Economic Association. Vol. N°5, Setiembre.

Slemrod, J. (2001) "A General Model of the Behavioral Response to Taxation". International Tax and Public Finance. Vol 8, 119-128.

Slemrod, J. y S. Yitzhaki. (2002). "Tax Avoidance, Evasion, and Administration," in A. Auerbach and M. Feldstein (eds.), Handbook of Public Economics, Vol. 3 (Amsterdam: NorthHolland), 1423-1470.

Virmani, A. (1989), "Indirect Tax Evasion and Production Efficiency", Journal of Public Economics, 39, 223-237.

Yaniv, G. (1988), "Withholding and non-withheld tax evasion", Journal of Public Economics 35(2), 183-204.

Yitzhaki, S. (1974), "A note on 'income tax evasion: a theoretical analysis", Journal of Public Economics 3(2), 201-202.

Yitzhaki, S. (1987), "On the excess burden of tax evasion", Public Finance Quarterly 15(2),123-137.