2-2017

VAT Treatment of Financial Institutions

Ismail BAYDUR
Singapore Management University, ismailb@smu.edu.sg

Fatih YILMAZ

VAT Treatment of the Financial Services: Implications for the Real Economy*

Ismail Baydur†
Singapore Management University

Fatih Yilmaz‡
Central Bank of the Republic of Turkey

February 15, 2017

Abtract: This paper studies the effects of exempt treatment of financial services under a VAT system. We develop a general equilibrium model with elastic labor supply, endogenous entry, and a banking sector. The banking sector provides loan services to producers and payment services to consumers. Our model display three key distortions under exempt treatment: (i) self-supply bias in the banking sector, (ii) consumption distortions, (iii) input distortions and tax cascading. Then, we calibrate our model to match the salient features of the tax system EU countries. A tax neutral policy regime switch from exempt treatment to full-taxation in loan services improves welfare about 4%. Shutting down the entry margin has even bigger welfare gains. The same policy exercise for payment services also implies welfare gains and these gains greater than zero rating of payment services.

*Fatih Yilmaz is grateful to Kenneth J. McKenzie for his support and guidance. We also benefited from comments of the seminar participants at the 2014 IIPF workshop in the Oxford University, and at the departmental seminars in University of Calgary, Ozyegin University, Istanbul Technical University, Central Bank of Turkey and ADA University. We particularly like to thank Katheryn N. Russ for sharing the MATLAB code, and Jevan Cherniwchan, Ali Shajarizadeh, and Trevor Tombe for their very valuable comments.
†Contact: ismailb@smu.edu.sg
‡Contact: fatih.yilmaz@tcmb.gov.tr

1 Introduction

Exempt treatment is the standard approach to taxing financial services in most of the countries with the value added tax (VAT) system.[1] Under this approach, financial institutions can neither charge VAT on their sales nor receive input credits for VAT paid on their input purchases. As such, the VAT is embedded in the price of exempt financial services. The resulting increase in the price of financial services creates several distortions in the economy. Three key distortions, in particular, have received most of the attention in the literature and the recent policy debate: (i) under-taxation of financial services to consumers; (ii) self-supply bias in the financial sector; and (iii) tax cascading in the real sector.[2] In this paper, we quantify the effects of exempt treatment on production and consumer welfare in a general equilibrium framework which features all of these three distortions.

We develop a representative household model with production and financial sectors. The financial sector is composed of banks that provide payment services to the representative household and loan services to the production firms. The banks combine labor and a taxable input to provide these services, but can only partially recover the VAT paid on their input purchases. The uncovered portion of the VAT is embedded in the price charged from the household and the production firms. We introduce VAT recovery rate, the covered portion of the VAT, as a tax policy instrument and incorporate this policy parameter to the banks' cost minimization problem. We fully characterize the pricing of the financial services in terms of the recovery rate and analyze the model's outcome under different tax regimes.

A more distinctive feature of our model is endogenous firm entry. While the existing studies focus only on the effects of a change in the tax policy regime at the intensive margin, we are able to study the effects of such a policy change at the extensive margin as well. Accounting for the extensive margin is especially important for evaluating the effects of policy changes on loan services. Most small and medium sized enterprises (SMEs) finance their investments largely through mainstream financial intermediaries such as banks, whereas large firms usually have better access to other means of financing, e.g. equity.[3]. This observation suggests that an increase in the cost

[1] For a cross-country policy discussion, see Gendron (2007) and Schenk (2009).

[2] The literature identifies another distortion that result from the exempt treatment of the financial sector: import bias - "an impediment to international competitiveness", which we omit here. See McKenzie and Firth (2011).

[3] See Denis and Mihov (2003) on bond issue and firms size and Houston and James (1996) and Johnson (1997) on

of financing due to switching to exempt treatment forces low productivity firms to exit the market and deter entry into the market.

We model endogenous firm entry by extending the framework in Melitz (2003). Although we extend the model in several dimensions, the model is still fairly tractable. In particular, we are able to show these distortions mentioned above analytically in our model and explain how they work in our model. First, due to complementarity in the production of banking services, banks also use their own resources in providing their payment services. However, under exempt treatment, they are not allowed to charge VAT on these inputs. Therefore, consumers pay a lower tax rate on their purchases of payment services, which distorts consumer's decisions. Second, the taxable input is relatively more expensive under exempt treatment because banks cannot recover the tax paid on these inputs under this regime. This relative price effect leads banks to employ more of their own resources and distorts resource allocation in the financial sector. Finally, the uncovered part of the VAT is embedded in the price of loan services. The embedded VAT not only increases the cost of financing for the production firms, but also creates tax cascading because production firms factor in this hidden VAT into their pricing decision.

We then calibrate our model to match the salient features of the tax system in the European Union (EU) countries and conduct several policy experiments. Our first finding is that the effects of exempt treatment on the extensive margin are sizable despite relatively small changes in cost of financing. In our benchmark calibration, where the VAT rate is 18%, the cost of the loan and payment services account for about 2% of total loans issued and 1.5% of total purchases mediated by banking services. Switching from exempt treatment to full taxation changes their shares by only 0.1 percentage points. While the costs of banking services marginally decreases after switching from exempt treatment to full taxation, the impact on the production side is sizable. Repealing exempt treatment for loan services increases the successful entry probability from 0.78 to 0.85 accompanied with 7.78% increase in the mass of firms in equilibrium.

Given the sizable impact on the production side, we evaluate the tax regime switch at different VAT rates. Our benchmark calibration implies that welfare increases at every tax rate under full recovery of loan services. The result is a reflection of the distortion on production sector. In our model, bank-to business (B2B) transactions are overtaxed under exempt treatment and moving

bank loans concentrating among small firms in the U.S.

from exempt treatment is welfare improving. We find the opposite result for payment services. Under full recovery of payment services, welfare is lower at every tax rate. The opposite result from payment services stems from the fact that consumers are under-taxed under exempt treatment.

We next turn to the analysis of tax revenue. Our calibrated model has a single peaked Laffer curve, which shows how total tax revenue changes with the VAT rate, at around 19% VAT rate. Moreover, as opposed to the welfare implications, switching to full taxation does not necessarily increase total tax revenue. We then performed the policy regime switch by keeping the tax revenue constant. We have four main findings in this analysis.

First, switching to full-taxation in loan services, does not increase the total tax revenue at every VAT rate, because the increase in the tax base is not necessarily large enough to compensate the tax revenue loss. Nonetheless, welfare improves after switching to full-taxation and the gains increase with the VAT rate reaching to a maximum of 4% gain.

Second, the gains from full-taxation are even larger, at a maximum of 5%, when we impose a lump-sum tax on profits so that the marginal productivity firm remains the same. This result suggests that full-taxation encourages entry of the low productive firms and misallocates capital from high productive firms to these low productive firms. A corrective tax is welfare enhancing in our model.

Third, even though full taxation of payment services reduces welfare at a given tax rate, gains from tax revenue is high enough to result in overall welfare gains over exempt treatment. Nonetheless, the welfare gains are smaller at around 2%. Fourth, when the tax on payment services is completely repealed, i.e. zero-rated, there are still welfare gains over exempt treatment, but less than the full-taxation case.

There appears to be no controversy regarding the taxation of B2B transactions, at least in principle. The clear consensus is that they should not be taxed and this can be accomplished in two ways: via full taxation treatment or by zero-rating B2B transactions. In our framework, these two methods are equivalent for B2B transactions. Our findings accord with the literature in that full-taxation of loan services is welfare enhancing. However, we also conclude that the gains from full-taxation is even bigger if there is a corrective tax on profits so that low productive firms are kept outside of the market and capital is allocated to high productive firms.

B2C transactions are not as straight forward and economists have different views on how to treat them. The views ranges from no taxation at all to taxation of both fee and margin based services. In our model, banks provide only fee based service to customers, i.e., payment services. Therefore, our finding supports the view of taxation of fee based services. This regime is also superior to zero-rating treatment where consumers do not pay any tax on payment services. Our result is a reflection of explicitly modeling the production payment services where there are real costs of providing these services, e.g. labor and input costs. A tax on these services reduces the excessive use of resources on the production of these services under exempt treatment.

The paper is organized as follows. The next section describes our model. We characterize and discuss the equilibrium in section 3. Section 4 calibrates the model. We perform po;icy analysis in Section 5. The last section concludes.

2 The Model

2.1 Overview

The model is set up in discrete time. There are five agents in the model: intermediate goods producers, a final good producer, banks, households, and government. The role of government is to collect income and value-added tax and transfer the receipts to households in a lump-sum fashion. Households maximize expected lifetime utility with a discount factor of β. In addition to transfers from the government, households earn labor and capital income, and receive firm profits.

There is an intermediate goods sector populated by monopolistically competitive firms each of which produces a differentiated good. A final good producer buys these differentiated products to combine them into a final consumption good. While the final good sector does not use any factors of production, the production technology in the intermediate goods sector uses both labor and capital. The capital market is facilitated by banks which accept deposits from households and issue loans to the intermediate good producers. In addition to loan services, banks also provide payment services to the households for their consumption good purchases.

Different from the final consumption good, the capital good is a homogeneous good produced in a competitive market according to a production technology using only labor. Henceforth, we use homogenous good and capital good interchangably. A distinctive feature of the banking activities is that banks hire labor and purchase inputs from the homogenous good market to produce banking services. These purchases of the banks are subject to VAT, but bank services can be tax exempt.

2.2 The Representative Household

Using prime ($'$) to denote next period variables, the representative household's problem can be written recursively as follows:

$$v(K) = \max_{C,N,K'} \quad \frac{1}{1-\eta}\left[C - \frac{N^{1+\psi}}{1+\psi}\right]^{1-\eta} + \beta v(K')$$

subject to the budget constraint:

$$(1+\lambda \bar{f}_s)(1+\tau)PC + (1-t_k\kappa)(K' - (1-\phi)K) = (1-t_w)N + r(1-t_k)K + LST \qquad (1)$$

The representative household derives utility from consumption, C, and dislikes labor, N. The utility function takes form in Greenwood et al. (1988) where the curvature parameter, ψ, is the inverse of the Frisch elasticity of labor supply. We choose labor to be the numéraire and normalize the wage rate to 1, which is subject to an income tax at a rate of t_w.

The representative household can also smooth consumption by adjusting his wealth, K, measured in terms of the capital good. We assume that the production of one unit of the capital good requires one unit of labor. This implies that the price of the capital good is also 1. All savings are deposited at the banks for a risk free rate of return, r, which are then loaned to intermediate good producers. The return to capital is subject to a tax, t_k, along with an investment allowance (e.g. investment subsidy), κ, which proxies the present value of tax depreciation deduction on capital investment, net of physical capital depreciation, ϕ. The consumption good is sold at the (relative) price P and is subject to a VAT at rate τ. We assume that λ fraction of the consumption good purchases are made through banking payment services.[4] The price paid on payment services is $\bar{f}_s$, which includes any VAT charged by the bank for its services. Finally, lump-sum transfers, LST, consist of profits from the intermediate goods sector and the total tax revenue. We explain them in more detail in the subsequent sections.

We focus on the steady state equilibrium. The first order condition for K' implies that the

[4] Lockwood (2011) endogenizes λ by inserting the disutility due to not using bank services.

steady state value of the real return on savings, r, is:

$$r = \frac{1 - t_k \kappa}{1 - t_k} \left(\frac{1}{\beta} - 1 + \phi \right). \tag{2}$$

Using the first order conditions for C and N, labor supply is obtained as follows:

$$N = \left[\frac{1 - t_w}{(1 + \lambda \bar{c}_s)(1 + \tau) P} \right]^{\frac{1}{\psi}} . \tag{3}$$

2.3 Final Good Production

A representative firm produces a final consumption good, Y, by combining intermediate goods, $y(\omega)$, according to a CES production function:

$$Y = \left(\int_{\omega \in \Omega} y(\omega)^{\frac{\sigma - 1}{\sigma}} \right)^{\frac{\sigma}{\sigma - 1}} . \tag{4}$$

The representative firm sets the final good price, P, competitively, i.e. P is equal to the marginal cost of production in equilibrium. To determine P, we obtain the marginal cost from the cost minimization problem of the final good producer below:

$$\min_{y(\omega) \in \Omega} \int_{\omega \in \Omega} p(\omega) y(\omega) \tag{5}$$

Solving the problem in (5) subject to equation (4) renders the standard Dixit-Stiglitz demand function for each variety:

$$y(\omega) = \left(\frac{p(\omega)}{P} \right)^{-\sigma} Y. \tag{6}$$

where P is the price index defined as follows:

$$P = \left(\int_{\omega \in \Omega} p(\omega)^{1 - \sigma} \right)^{\frac{1}{1 - \sigma}} \tag{7}$$

2.4 Banking Sector

Banks produce two streams of services in a perfectly competitive environment: loan services for firms and payment services for consumption good purchases. We assume that there are no economies of scope in the production of loan and payment services, and each stream operates independently of each other.

In practice, along with their own resources, banks use outside resources purchases of which may be subject to VAT. There is a wide variation in policy practices across countries, but, in most instances, purchases of the "arranging for" services by financial institutions are subject to VAT. These services include checking, evaluating or authorizing credits, monitoring payment records, credit management services, credit card promotion, market research, document processing and preparation, research, analysis and assessment services, payroll and IT services. When a bank provides a service that is exempt from VAT, it is unable to recover the resulting VAT embedded in the price of these inputs.

We model the banking services in the next two subsections in line with these observations. To capture the essence of the taxation of financial services in practice, we assume banks employ labor and use the homogenous good as input in the production of financial services. Purchases of the homogenous good are subject to VAT, but banks can only partially cover the VAT paid on these inputs.

2.4.1 Loan Services

A bank incurs monitoring and administrative costs to extend a loan to an intermediate good producer. To produce l units of loan services, the bank uses labor, n_l, and the homogenous good as input, y_l, according to a Cobb-Douglas production function:

$$l = A_l n_l^{1-\zeta} y_l^{\zeta} \tag{8}$$

Unlike capital, banks exhaust the homogenous input good completely in the production process. We also assume that households can freely move between the intermediate goods and banking sectors and are paid at the same rate regardless of the sector they are employed in. This latter assumption implies that the prices of both of the bank inputs are the same, but the purchases of

y_l are subject to a VAT at rate τ. Therefore, by using its own resources, a bank can avoid paying VAT on its inputs and this property of the production function allows us to examine the effects of changes in the tax regulations on the resource allocation in the banking sector.

When extending financial services to intermediate good producers, banks can charge VAT for ρ_l fraction of their services and fully recover VAT paid on y_l, while $(1-\rho_l)$ fraction of their services are under exempt treatment, i.e. banks cannot charge VAT on these services. ρ_l is a policy tool in the tax system and determined by the government. Accordingly, we can write the cost minimization problem for producing one unit of loan services as follows:

$$C_l(n_l, y_l) = \min_{n_l, y_l} \rho_l\,(n_l + y_l) + (1-\rho_l)\,(n_l + (1+\tau)y_l) \tag{9}$$

subject to (8) with $m = 1$. Note that, for the part of the services that are under exempt treatment, the effective cost to the bank of using taxable input is $(1+\tau)$.[5] First order conditions yield input demand functions as follows:

$$n_l^* = \frac{1}{A_l}\left(\frac{(1-\zeta)R_l}{\zeta}\right)^{\zeta}, \tag{10}$$

$$y_l^* = \frac{1}{A_l}\left(\frac{(1-\zeta)R_l}{\zeta}\right)^{\zeta-1}, \tag{11}$$

where we define $R_l = \rho_l + (1-\rho_l)(1+\tau)$. The unit cost of loan services, f_l, can be calculated as:

$$f_l = C_l(n_l^*, y_l^*) = \frac{R_l^{\zeta}}{A_l Z}, \tag{12}$$

where $Z = \zeta^{\zeta}(1-\zeta)^{1-\zeta}$. Furthermore, a bank must produce $(1 + aK)$ units of monitoring and administrative services to issue a loan of size K. The constant term, which we normalize to 1, captures the fixed costs associated with the loan services. In practice, these fixed costs might include any paper work or a standard procedure in evaluating the loan application. The second term captures the variable cost of the loan services extended. The variable costs might be associated with following up for the timely payments from the clients or additional monitoring activities to reveal more information about the default risk of the borrower. In our model, banks charge the

[5] See (UK source) for an illustration of how VAT for the banking sector is calculated in practice.

fixed component, f_l, as a one time payment, which we refer to as loan application fee, and the variable cost part, af_l, is embedded into the loan interest rate.

Next, we specify the loan interest rate, r_l. Intermediate goods producers are subject to a uniform exogenous default risk, δ. Without loss of generality, we assume that the loans are fully collateralized, i.e., defaulting firms exit the market and return the loan principal, but are unable to pay the loan interest. Nevertheless, the bank still has to pay interest to its depositors, at the rate of r. In equilibrium, banks make zero profit. Hence, expected return from a loan of size K is equal to zero: $((1-\delta)r_l - r - af_l)K = 0$. After rearranging,

$$r_l = \frac{r + af_l}{1-\delta}. \tag{13}$$

This equation illustrates the practical problem experienced in taxing financial institutions. The price charged by the bank is a combination of the risk premium due to the default probability, δ, and the variable cost of the loan services, af_l. A VAT imposed directly on r_l would imply taxing the risk premium. Ideally, a tax authority would want to subtract the risk premium, because it is a cost component. However, δ is not verifiable in practice.

Note also that R_l is decreasing in ρ_l. This result has three implications which are basis for the distortions we study in the policy analysis below. First, the ratio of equations (10) and (11) imply that,

$$\frac{n_l^*}{y_l^*} = \frac{1-\zeta}{\zeta} R_l. \tag{14}$$

As ρ_l increases, e.g., moving from full taxation (ρ_l=1) to exempt treatment ($\rho_l = 0$), makes banks employ more labor relative to taxable input. This relationship establishes one of the distortions we want to show in this paper: exempt treatment leads to "self-supply bias" in the banking sector. From the production function in equation (8), this price distortion leads to inefficiencies in the banking sector. Second, moving from full taxation to exempt treatment raises the fixed loan application fee, f_l. Such a policy change increases the borrowing cost of intermediate good producers and distorts resource allocation in that sector. Third, under exempt treatment, VAT paid by the banks is embedded in r_l. When selling their products to the final consumption good producer, intermediate good producers factor in these costs into their pricing decisions. Consumers pay VAT on their purchases which already includes a VAT paid by the banks. Through this channel, we obtain a tax cascading effect in our model.

2.4.2 Payment Services

Payment services are provided to households for their purchases of consumption good. To facilitate the purchase of the consumption good, a bank has to produce s units of payment services. Similar to loan services, these services are produced using labor and inputs from the homogenous good sector according to a Cobb-Douglas production function:

$$s = A_s n_s^{1-\gamma} y_s^{\gamma} \tag{15}$$

With a slight change in notation, cost minimization problem is as follows:

$$C_s(n_s, y_s) = \min_{n_s, y_s} \rho_s \left(n_s + y_s\right) + (1 - \rho_s)\left(n_s + (1+\tau) y_s\right) \tag{16}$$

subject to (15). Unlike the loan services, there is no heterogeneity on the consumer side. As long as the number of transactions is a constant fraction of the total consumption, assuming fixed or variable cost of payment services imply the same pricing rule for banks.[6] Accordingly, we assume that s units of payment services must be produced whenever the representative household purchases a consumption good and normalize s to 1. First order conditions from the problem in (16) yield input demand functions as follows:

$$n_s^* = \frac{1}{A_s}\left(\frac{(1-\gamma)R_s}{\gamma}\right)^{\gamma}, \tag{17}$$

$$y_s^* = \frac{1}{A_s}\left(\frac{(1-\gamma)R_s}{\gamma}\right)^{\gamma-1}, \tag{18}$$

[6] We elaborate more on this issue in calibration section. In practice, fixed card fees are negligible when compared to merchant service fees.

where $R_s = \rho_s + (1-\rho_s)(1+\tau)$. The fixed cost of processing any amount of final consumption good purchase can be calculated as:

$$\tilde{f}_s = \quad C_s(n_s^*, y_s^*) = \tfrac{R_s^\gamma}{A_s\Gamma}, \tag{19}$$

where $\Gamma = \gamma^\gamma(1-\gamma)^{1-\gamma}$. Different than the fixed cost to the bank, the cost to the consumer, $\bar{f}_s$, includes VAT for the recovered part of the payment services, i.e. $f_s + \rho_s\tau(n_s + y_s)$. Inserting input demand function from equation (17) yields:

$$\bar{f}_s = f_s\left[1 + \rho_s\tau\left(1 - \gamma + \frac{\gamma}{R_s}\right)\right]. \tag{20}$$

Note that, under full-recovery tax regime, $\rho_s = 1$, the price paid by the consumer inclusive of VAT is equal to $\frac{(1+\tau)}{A_s\Gamma}$. Under exempt treatment, $\rho_s = 0$, the price paid by the consumer is equal to $\frac{(1+\tau)^\gamma}{A_s\Gamma}$, which is less than the price under full-recovery as long as $\gamma < 1$. In other words, consumers are under-taxed under exempt treatment in our model. This feature of the model is consistent with the empirical literature and we study its welfare implications in the policy analysis section.

In general, $\bar{f}_s$ is increasing in ρ_s whenever $0 < \gamma < 1$. To understand the intuition, it is helpful to re-write $\bar{f}_s$ in terms of the input demand functions:

$$\bar{f}_s = \quad (1+\rho\tau)n_s + (1+\tau)y_s.$$

Taking the ratio of input demand functions in equations (17) and (18) and using the production function for payment services in the equation above, we obtain:

$$\bar{f}_s = \quad \tfrac{1}{A_s\Gamma}\left((1-\gamma)(1+\rho_s\tau)R_s^\gamma + \gamma(1+\tau)R_s^{\gamma-1}\right),$$

Taking the derivative of the above equation with respect to ρ_s and re-arranging, we obtain the following expression:

$$\frac{d\bar{f}_s}{d\rho_s} = \quad \frac{1}{A_s\Gamma}(1-\gamma)\tau R_s{}^\gamma\left(1 - \gamma\frac{1+\rho_s\tau}{R_s} + \gamma\frac{1+\tau}{R_s^2}\right) \tag{21}$$

There are three terms in parenthesis that determine the sign of this expression. The first term is equal to one and captures the direct effect of a change in ρ_s. As ρ_s increases, the total tax burden on consumers from the labor input also increases. The second and the third terms capture the relative price effect on the firm's cost minimization problem. As the recovery rate gets closer to one, banks can recover a larger portion of the VAT paid on taxable inputs. Cost minimization requires banks to use relatively less labor input and more taxable input in their production process. On one hand, this relative price effect reduces the tax burden stemming from the use of labor input. This effect is captured by the negative sign in front of the second term. On the other hand, the relative price effect raises the tax burden on consumers stemming from the use of taxable inputs. Hence, there is a positive sign in front of the third term above. Overall, the sign of the equation in (21) is positive as long as $\gamma < 1$.

2.5 Intermediate Goods Producers

This part of the model extends the standard Melitz (2003) model by incorporating capital into the production process. Each firm produces a differentiated good, denoted by φ, according to a Cobb-Douglas production function:[7]

$$y = \varphi n^{\alpha} k^{1-\alpha}. \tag{22}$$

φ is drawn upon entry from a fixed distribution with c.d.f. $H(\varphi)$ after paying a (real) fixed entry cost, f_e. φ, also a measure of productivity, remains the same until the firm is exogeneously destroyed by probability δ.

Because the market is monopolistically competitive, firms are also price setters. To characterize their optimal decisions, let us first derive the unit marginal cost of production from the firm's cost minimization problem:

$$\min_{n,k} n + r_l k \tag{23}$$

[7]To save on notation, we suppress the subscripts for firms. As we show below, all the firm decision can be summarized by φ.

subject to the production function in (22). Input demand functions are calculated as follows:

$$n* = \left(\frac{\alpha}{1-\alpha} r_l\right)^{1-\alpha} \frac{y}{\varphi} \tag{24}$$

$$k* = \left(\frac{\alpha}{1-\alpha} r_l\right)^{-\alpha} \frac{y}{\varphi} \tag{25}$$

Plugging back into (23) and taking derivative with respect to y yields the marginal cost of production as follows:

$$MC(\varphi) = \frac{r_l^{1-\alpha}}{B\varphi}, \quad \text{where } B = \alpha^{\alpha}(1-\alpha)^{1-\alpha}. \tag{26}$$

To find the optimal price charged by a firm with productivity φ, let us solve the profit maximization problem after inserting the demand equation in (6):

$$\max_p \pi(\varphi) = p^{1-\sigma} P^{\sigma} Y - MC(\varphi) p^{\sigma} P^{\sigma} Y - f_l \tag{27}$$

The last term is the loan application fee paid to the banks as financing cost.[8] The first order condition with respect to p_i renders:

$$p(\varphi) = \frac{\sigma}{\sigma - 1} MC(\varphi) \tag{28}$$

2.6 Equilibrium

To close the model, we impose a free entry condition: the expected discounted value of a potential entrant firm is equal to zero, i.e.,

$$\int \frac{\pi(\varphi)}{1-\beta(1-\delta)} dH(\varphi) = f_e. \tag{29}$$

where integration is over the domain of φ.

Assuming that f_m is sufficiently large, profits are negative for some of the low productive firms. Hence, there exists a reservation productivity, $\varphi*$ such that, in equilibrium, only firms with $\varphi \geq \varphi*$ produces. We solve for the equilibrium value of $\varphi*$ in the next section using this reservation

[8] Without loss of generality, we assume that the loan application fee internally financed and there is no interest charge on this component.

productivity property and the free entry condition.

Free entry condition also implies that there is a constant mass of firms, g, that is endogenously determined in equilibrium. In particular, the stationary condition requires that the mass of successful entrants are equal to the mass of exiting firms: $(1 - H(\varphi *))g_e = \delta g$.

Together with market clearing conditions, we are now in a position to define steady state equilibrium:

[definition of stationary equilibrium]

3 Characterization of Equilibrium

3.1 Equilibrium Reservation Productivity

The existence of a reservation productivity has two implications. First, the discounted value of a firm is $\pi(\varphi)/(1 - \beta(1 - \delta))$, if $\varphi \geq \varphi^*$ and zero otherwise. Hence, the expected discounted value of a potential entrant is the product of the probability of successful entry, $1 - H(\varphi^*)$, and the average discounted value of successful firms:

$$\bar{\pi} = \int_{\varphi^*}^{\infty} \frac{\pi(\varphi)}{1 - \beta(1 - \delta)} dH(\varphi|\varphi \geq \varphi^*). \tag{30}$$

Then, the free entry condition (FE) implies the following condition on $\bar{\pi}$:

$$\bar{\pi} = \frac{(1 - \beta(1 - \delta))f_e}{(1 - H(\varphi^*))} \tag{31}$$

Second, operating profits of a firm with reservation productivity makes zero profit: $\pi(\varphi^*) = 0$. This zero profit condition (ZP) implies that the revenue for this firm is $r(\varphi^*) = \sigma f_l$. Moreover, the revenue function is proportional to $\varphi^{\sigma-1}$. Using this latter property, we can write the average operating profits as follows:

$$\bar{\pi} = \left[\frac{\bar{\varphi}^{\sigma-1}}{\varphi^{*\sigma-1}} - 1 \right] f_l \tag{32}$$

where we define $\bar{\varphi}$ as the average productivity:

$$\bar{\varphi} = \left(\int_{\varphi^*}^{\infty} \varphi^{\sigma-1} dH(\varphi|\varphi \geq \varphi^*) \right)^{\frac{1}{\sigma-1}}. \tag{33}$$

At this point, we assume that φ is distributed according to a Pareto distribution with c.d.f. $H(\varphi) = 1 - \varphi^{-\theta}$. This assumption implies the following linear relationship between φ^* and $\bar{\varphi}$:

$$\bar{\varphi} = \left(\frac{\theta}{\theta - \sigma + 1}\right)^{\frac{1}{\sigma-1}} \varphi^*. \tag{34}$$

Plugging this back into ZP equation in (32), we can write the average operating profits as:

$$\bar{\pi} = \frac{\sigma - 1}{\theta - \sigma + 1} f_l. \tag{35}$$

ZP and FE conditions together with Pareto distribution assumption on productivity deliver a closed form solution for φ^*:

$$\varphi^* = \left(\frac{f_l(\sigma - 1)}{f_e(\theta - \sigma + 1)(1 - \beta(1 - \delta))}\right)^{\frac{1}{\theta}} \tag{36}$$

We depict the determination of equilibrium value of φ^* in Figure (1). Note that FE is always an

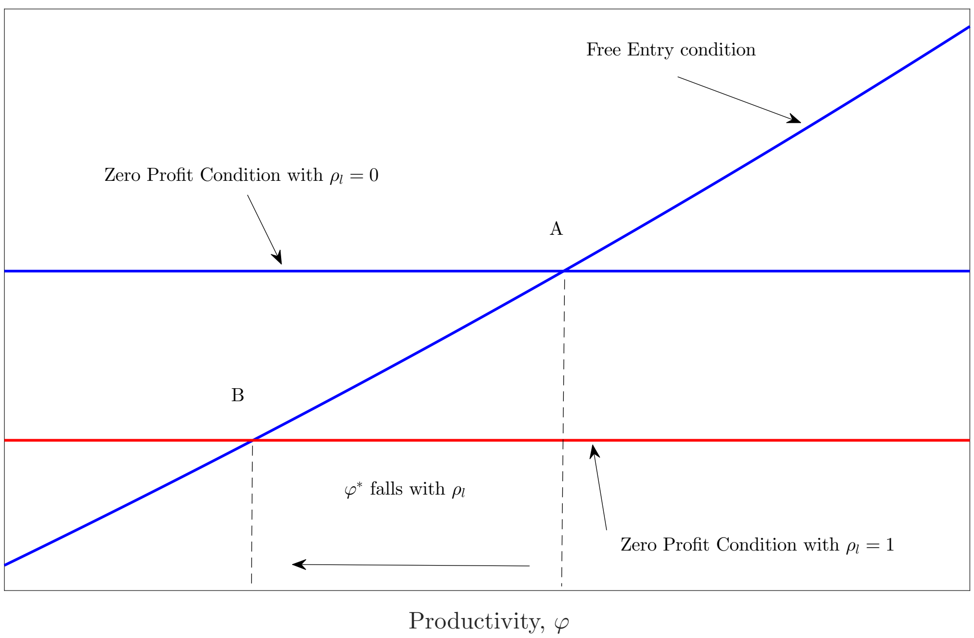

Figure 1: Determination of Productivity Threshold

increasing function of φ, but ZP is independent of φ^* due to the Pareto distribution assumption. Repealing exempt treatment and moving towards full taxation decreases the fixed cost of obtaining a loan, f_l, and leads to a downward shift in ZP line. This shift moves the equilibrium from point A to point B so that φ is lower. In other words, lower cost of capital makes low productive firms enter into the market. This effect on productivity threshold is one channel how tax policy can affect the extensive margin in the production sector. In the next section, we analyze the effect of the change in tax policy on the other equilibrium outcomes, in particular to the mass of firms, g, which is another extensive margin variable in the model.

3.2 Aggregation and Market Clearing

Note that monopolistic competition framework implies that the individual decisions of intermediate good producers are proportional to the functions of φ. This feature renders rather simple expressions for the aggregate variables.

Using the definition for average productivity in 33 and integrating over the optimal price, $p(\varphi)$, over all the operating firms, the price index, P, is obtained as follows:

$$P = [(1-\delta)g]^{\frac{1}{1-\sigma}} \frac{\sigma}{\sigma - 1} \frac{r_l^{1-\alpha}}{B\bar{\varphi}} \tag{37}$$

In the final goods markets, total revenue is equal to the total spending on inputs:

$$PY = \int_{\varphi^*}^{\infty} p(\varphi) y(\varphi) dH(\varphi | \varphi \geq \varphi^*). \tag{38}$$

Evaluating the integral over φ on the right hand side and using the expression for φ^* in equation (36) yields:[9]

$$PY = (1-\delta) g \sigma f_l \frac{\theta}{\theta - \sigma + 1}. \tag{39}$$

Note that P and Y can be evaluated once g is known. We calculate equilibrium value of g from the labor market clearing condition. Total demand for labor includes labor demand in the intermediate goods sector, banking sector, and the homogenous good sector. The labor demand in the homogenous good sector is equal to the investment expenditures, banking sector input, and

[9] This step uses Pareto distribution assumption.

the total fixed entry cost. First, total demand for labor in the intermediate goods sector is:

$$N_\varphi = (1-\delta)g\int_{\varphi^*}^{\infty} n(\varphi)dH(\varphi|\varphi \geq \varphi*) = \frac{\sigma-1}{\sigma}\alpha PY \tag{40}$$

Interpretation of this labor demand function follows straightforward Cobb-Douglas properties: α fraction of the total revenue after profits are paid to the labor. Similarly, total demand for capital is:

$$K = \quad g\int_{\varphi*}^{\infty} k(\varphi)dH(\varphi|\varphi \geq \varphi*) \quad = \frac{\sigma-1}{\sigma}\frac{(1-\alpha)Y}{(1-\delta)r_l}. \tag{41}$$

Total capital demand has an interpretation similar to the labor demand function in the intermediate goods sector. At the steady state equilibrium, total investment is $I = \phi K$, and this value adds to the total labor market demand.

In the banking sector, the production of one unit of homogenous good requires one unit of labor. Therefore, the labor demand due to loan services is the sum of these factors of production: $(n_l + y_l)(g + aK)$, or $f_l(g + aK)$ from the definition of the unit cost of loan services. Similarly, the labor demand due to payment services is $(n_s + y_s)\lambda(1+\tau)PY$, which is equal to $f_s\lambda(1+\tau)PY$. Note that we already used the goods market clearing condition here: $C = Y$. Finally, total entry cost incurred is equal to $g_e f_e$.

Setting equation (**??**) from consumer's problem equal to total labor demand gives us the mass of the firms in the equilibrium. Inspection of the total labor demand reveals that labor demand is linear in g, but the labor supply is proportional to $g^{\frac{1}{\psi(1-\sigma)}}$. We can write the labor market equilibrium in a compact form as follows:

$$N^s = N^d \rightarrow c_1(f_l, f_s)g^{\frac{1}{\psi(\sigma-1)}} = c_2(f_l, f_s)g, \tag{42}$$

where c_1 and c_2 are increasing functions of f_l and f_s. We plotted the labor market equilibrium in Figure 2. Note that $g = 0$ trivially satisfies this equation. For the interior solution, denoted by A, stability of equilibrium requires that labor supply function is flatter than the labor demand function at the intersection point. This requires that $\sigma > 1/\psi + 1$. In words, the elasticity of the substitution between the differentiated products must be sufficiently greater than the elasticity of

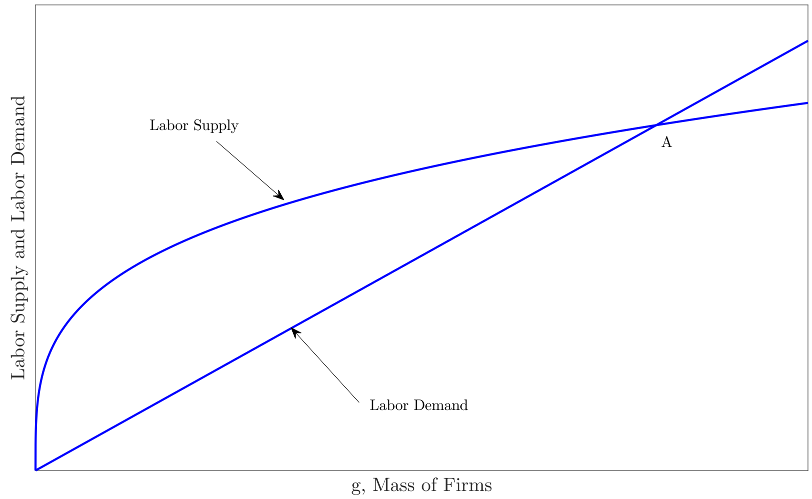

Figure 2: Determination of the Mass of Firms

labor supply.[10] We can analyze the effects of a tax regime change in the banking sector on the

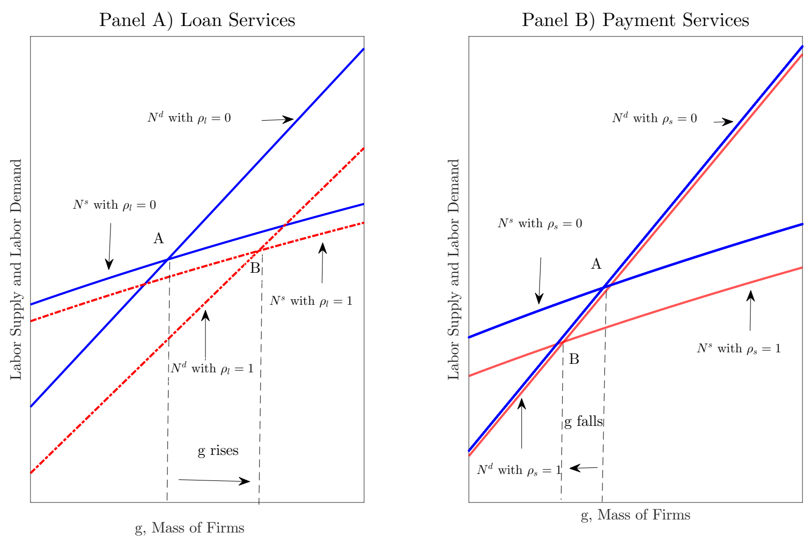

Figure 3: Equilibirum Mass of Firms after Switching from Exempt Treatment ($\rho_{l,s} = 0$) to Full Taxation ($\rho_{l,s} = 1$)

mass of firms. For example, switching from exempt treatment to full taxation in the loan services lowers f_l, which in turn lowers both the labor demand and supply. We show the effect of this policy change in panel A of Figure 3. Note that the effect on labor demand is larger and the mass of firms

[10]The standard Melitz(2003) model is the limiting case when $\psi \to \infty$. In that case, the labor supply curve is horizontal.

is larger at the new equilibrium. This result holds in general as long as $\sigma > 1/\psi + 1$. Recall that, a decrease in f_l also lower the productivity threshold. Hence, the total effect on the extensive margin of switching to full taxation in loan services (B2B services) is always positive.

A similar tax regime switch in the payment services has no effect on productivity threshold, but decreases both the labor demand and supply. We show the effect of this policy change in panel B of Figure 3. In the figure, there is a smaller mass of firms at the new equilibrium, but, in general, the effect on equilibrium g is ambiguous and depends on the elasticity of labor supply.[11] Recall that consumers are under-taxed under exempt treatment and switching to full taxation increases the tax paid by the consumers. In response to the increased tax burden on consumption, the representative consumer substitutes from consumption to leisure. The decrease in labor supply has a negative impact on the mass of firms in equilibrium.

It is of our interest to know the magnitude of these effects on the extensive margin and consumer welfare and how they interact with the other tax policy instruments. In the next section, we quantify these effects using a calibrated version of the model.

4 Calibration

Table 1 shows the calibrated parameter values. We set a period to be one year and the discount factor, β, equal to 0.960, which implies an annual interest of 4.2%. Following the Real Business Cycle (RBC) literature, we set the physical capital depreciation rate, ϕ, equal to 0.100. Due to the existence of homogenous good sector, the labor share of production in the differentiated goods sector is different than the curvature parameter of the production function, α. Using plant-level data available from the Longitudinal Research Database, Cooper, Haltiwanger, and Willis (2004) estimate $\alpha = 0.640$ and we use this value in our calibration. This value implies that the share of labor in total production is 71.84%, which is consistent with the estimates in Gomme and Ruppert (2007). Motivated by the estimates in Feenstra (1994), we set the elasticity of substitution across differentiated goods, σ, equal to 8.000. Note that firm revenues are also Pareto distributed with the shape parameter $\theta/(\sigma - 1)$. We impose $\theta = \sigma$ so that the resulting distribution of firm revenue

[11] In the extreme case where labor supply is horizontal as in the original Melitz model, labor supply would be unaffected and g would increase in equilibrium.

is consistent with the estimates in Helpman et al. (2004).[12]

Table 1: Calibrated Model Parameters

Parameter		Target	Value
		Production and Consumer Preferences	
β:	Discount Factor	Annual interest rate of 4.2%	0.960
α:	Labor Share (Differentiated Goods)	Cooper, Haltiwanger, and Willis (2004)	0.640
ϕ:	Depreciation Rate	RBC literature	0.100
σ:	Elasticity of Substitution (Differentiated Goods)	Feenstra (1994)	8.000
θ:	Pareto distribution shape parameter	Helpman et al. (2004)	8.000
ψ:	Frisch elasticity of labor supply	$1/\psi$=2.000	0.500
f_e:	Fixed entry cost	Average firm size of 20	5.859
		Tax System	
τ:	VAT rate	OECD data (Europe)	0.180
t_w:	Payroll tax rate	OECD data (Europe)	0.227
t_k:	Capital income tax rate	OECD data (Europe)	0.220
κ:	Investment allowance	Bettendorf et al. (2009)	0.750
		Banking Sector	
ζ, γ:	Labor Share (loan & payment services)	Huizinga (2002)	0.450
ρ_l, ρ_s:	VAT recovery rate	Büttner and Erbe (2012)	0.333
δ:	Default probability	Russ and Valderrama (2012)	0.050
λ:	Share of PC using payment services	Bagnall et al. (2014)	0.440
a:	Variable cost loan services	Interest rate spread: 2%	0.124
A_s:	Productivity (payment services)	Merchant service fee: 1.5%	147.577
A_l:	Productivity (loan services)	Survival probability, ${\varphi^*}^{-\theta} = 0.8$	22.744

We set income and value-added tax rates in the model to the median statutory tax rates in OECD countries. Accordingly, we set the VAT rate, τ, equal to 0.180. The payroll tax rate, t_w, and the capital income tax rate, t_k, are equal to 0.227 and 0.220, respectively. There is no direct measure of the present value of investment allowance, κ. Following Bettendorf et al. (2009), we set $\kappa = 0.750$, which is the average value estimated for EU countries.

Our policy experiments are sensitive to the value of the Frisch elasticity of labor supply, because any change in τ uniformly affects the price of the differentiated goods. If labor supply is highly inelastic, an increase in τ leads to ever increasing tax revenues. If, on the other hand, the labor supply is highly elastic, an increase in τ leads to ever decreasing total tax revenue. We set ψ equal to 0.500 so that the Frisch elasticity of labor supply is equal to 2. This value is at the lower bound

[12]Note that, for some industries, the implied shape parameter for the revenue distribution is less than 1, which violates our restriction for finite firm size: $\theta \geq (\sigma - 1)$. Nonetheless, our parameter choice implies that the shape parameter is close to one.

of existing macroeconomic estimates for labor elasticity and implies a peak in the Laffer curve at around 19% VAT rate. In the online appendix, we perform a sensitivity analysis for this parameter value.

Following Russ and Valderrama (2012), we use a middle ground value for the loan default rate, δ, and set it equal to 0.050. The other banking sector parameters in the model are less documented. Huizinga (2002) reports that the share of employee costs in total operating costs is 55% in the EU countries. We choose the curvature parameters in the production functions of the loan and payment services, ζ and γ, to match this figure. In the absence of separate estimates of this variable for B2B and B2C transactions, we set both ζ and γ equal to 0.450. Büttner and Erbe (2012) estimates that about one-third of the total financial services in Germany are taxable. Unfortunately, this study also does not make a distinction between loan and payment services. Hence, we set both ρ_l and ρ_s equal to 0.333.

Using cross-country dairy surveys, Bagnall et al. (2014) calculates that the share of non-cash transactions in consumer spending shows great cross-country variation, ranging from 18% in Germany to 54% in the U.S. These transactions are mediated by banks, and in most cases either by a debit or a credit card. We set the share of transactions that uses payment services in the model, λ, equal to 0.440. This value is reported for France and corresponds to the median value in Bagnall et al. (2014).

According to a report by EU Commission in 2006, merchant service charges also show great variation across European countries. Depending on the type of the card used in transaction (credit vs. debit), location (domestic vs. international), firm size and industry, the costs borne by the merchants are from 0.50% for a domestic debit card transaction to 3.20% for an international credit card transaction. Although merchant service charges are sizable, the cases where consumers pay a transaction fee is rare in European countries, and if a consumer ever pays a transaction fee, it is typically less than 1%. In addition, card issuance and annual fee charges are small, on average 24 Euros per credit card. Nonetheless, one would expect that sellers would reflect their share of the payment service fee on the price, at least partially. Based on this information, we choose the productivity parameter in payment services, A_s, to reflect the merchant service fees reported in this study. Given that debit card fees are smaller and most of the transactions use a debit card, we choose a conservative target at 1.5% and equate payment services fee in the model, $\bar{f}_s$, to this number. The implied value for A_s is 147.577.

There are three more parameters to be calibrated in the model: the scale parameter and share of variable costs in loan services production, A_l and a, and the fixed entry cost, f_e. Recall that successful entry in this model depends on drawing a firm productivity higher than a threshold. To draw its productivity, a firm has to pay an entry cost. An interpretation of this entry cost is that a potential entrant firm incurs these costs implicitly by performing an initial investment, a pilot production, or conducting a marketing research. By doing so, the firm learns about its productivity. Along these lines, we interpret the successful entry probability in the model, $\varphi^{*-\theta}$, as the survival probability of an entrant firm in the first year. Using data from OECD countries for the period from 1990 to 2000, Bartelsman et al. (2003) reports that 75% to 85% of the new firms survive in their first year. Given the value of θ, we target $\varphi^* = 1.028$ in equilibrium so that 80% of the firms decide to produce after drawing their productivity level. Note that φ^* depends on the ratio of f_l and f_e and the value of f_l depends on A_l. For a given value of f_e, there is a unique value of A_l that achieves our targeted productivity cut-off level and we calibrate A_l to this value.

While the fixed loan application fee affects the productivity threshold, the variable cost component directly enters the loan interest rate and increases the spread between the loan and deposit interest rate, $(r_l - r)$. For the given value of default probability, which captures the risk premium in the spread, we calibrate the share of variable costs, a, so that the interest rate spread is 2% in the model. Our targeted value corresponds to average interest rate spread in EU countries in the first half of 2000s and before the Great Recession.

Before explaining the calibration of f_e, it is helpful to report certain equilibrium ratios implied by our calibration so far. The share of the value of inputs used by the banks in total output is 3.64% and the banking sector employs 2.93% of the total labor supply. These numbers are reasonable given that the entire financial sector produces roughly 3.5% of the output in Europe.[13] Total loan application fees collected is only 0.85% of the total capital loaned out to the intermediate good producers, which accords with the estimates for the U.S. banks provided by DeYoung and Rice (2006). Income shares of the factors of production are also at reasonable ranges. The shares of labor and capital income in total output before any taxes is 71.84% and 24.41%, respectively. The remaining 3.74% of the output is attributable to firm profit.

[13] See Huzinga (2002).

Note that these ratios are independent of the actual value of f_e and hold as long as our targeted productivity cut-off level is satisfied. In other words, the exact value of f_e only determines the scale of the economy. We choose f_e to target an average size of 20 for intermediate good producers. Taking N units of labor supply equivalent to one worker, the average firm size is given by $n(\bar{\varphi})/N_\varphi$. From equation (40), this targeted value implies that equilibrium mass of firms, g, must be such that $g(1-\delta) = 0.05$. This equilibrium target requires f_e to be equal to 5.859. For this value of f_e, the implied values of A_l and a are 22.744 and 0.124, respectively.

5 Policy Analysis

In this section, we analyze the effects of switching from exempt treatment, i.e., the recovery rate is zero, to full taxation, i.e., the recovery rate is one. We perform our analysis separately for loan and payment services. We also compare our analysis to two other cases. For the loan services, we compare the results under exempt treatment to those under the full taxation with a lump-sum profit tax. The lump-sum profit tax is such that the equilibrium productivity threshold remains constant despite the tax regime change. By doing so, we isolate the effects of the newly entering low productivity firms. For payment services, we compare our results to a zero-rated tax exemption regime where consumers do not pay tax on payment services.

Table 2: Percentage Changes in Equilibrium Variables after Repealing Exempt Treatment in Loan Services ($\rho_l = 0$): Tax rate is held constant at $\tau = 0.18$.

Variable (Change in %)	$\rho_l = 1$	Profit Tax
Entry Probability ($\varphi^{*-\theta}$)	7.448	- - -
Average Productivity ($\bar{\varphi}$)	−0.931	- - -
Mass of firms (g)	7.989	0.360
Employment (Differentiated Good)	0.541	0.360
Extensive Margin (φ)	0.932	- - -
Extensive Margin (g)	7.648	0.360
Intensive Margin	−8.038	- - -
Total Tax Revenue	0.049	0.091
Payroll Tax	0.791	0.474
Capital Tax	1.056	0.875
VAT	−0.931	−0.460
Output	0.624	0.443
Welfare (logarithm of instantaneous utility)	0.786	0.528

Table 3: Percentage in Equilibrium after Repealing Exempt Treatment in Payment Services ($\rho_s = 0$): Tax rate is held constant at $\tau = 0.18$.

Variable (Change in %)	$\rho_s = 1$	Zero-Rated
Entry Probability ($\varphi^{*-\theta}$)	- - -	- - -
Average Productivity ($\bar{\varphi}$)	- - -	- - -
Mass of firms (g)	−0.166	0.132
Employment (Differentiated Good)	−0.166	0.132
Extensive Margin (φ)	- - -	- - -
Extensive Margin (g)	−0.165	0.132
Intensive Margin	−0.001	0.000
Total Tax Revenue	0.013	0.004
Payroll Tax	−0.168	0.129
Capital Tax	−0.166	0.132
VAT	0.239	−0.152
Output	−0.168	0.130
Welfare (instantaneous utility)	−0.152	0.125

We first keep the VAT rate at 18%, which we used in our calibration, and change the recovery rate from zero to one in a separate analysis for loan and payment services . The first column in Table 2 shows the percentage changes in various equilibrium values after switching to full taxation in loan payment services. Under full-taxation, the entry probability increases by around 7.5% relative to its equilibrium value under exempt treatment. The new entrants at the lower tail of the distribution lowers the average by productivity by about 1%. The change in the mass of firms is also sizable and increases by about 10% after repealing the exempt treatment. The effect of tax policy regime on total employment in the differentiated goods sector is half a percent increase and this total effect is a combined effect of extensive and intensive margins. We find that existing firms lower their employment by 8%, but this decrease is largely balanced with the increase in the mass of firms which are at least as productive as the marginal firm under exempt treatment. In addition to that, the newly entering firms which would not be profitable under exempt treatment increases the employment by 1%. The increase in employment results in a 0.6% increase in total output inclusive of the homogenous goods sector, and about a 0.8% increase in welfare.

To assess the impact of endogenous entry, we impose lump-sum tax on the profit of the firms in a way that keeps the productivity threshold unchanged while switching to full-taxation regime. We report the results in the second column of Table 2.[14] By construction, all the changes come

[14]We added the profit tax to the total of VAT revenue.

from the change in the mass of firms. Compared to the first column, the response of employment, output, and welfare are muted, accounting for roughly two-thirds of the increase in these variables without the profit tax. Nonetheless, the impact on total tax revenue is bigger, which we evaluate further below.

In the first column of Table 3, we report the results from switching to full-taxation in payment services. Compared to regime switch in loan services, the response of the equilibrium outcomes are much smaller. First, the tax regime switch does not change the productivity threshold. The price of payment services decreases, but its impact on the intensive margin negligible. Overall, under full-taxation, employment and output decrease by 0.16% and imply a decrease in welfare of similar magnitude. In the second column, we report the results under a zero-rated regime, under which neither the consumers nor the banks pay a tax on payment services. Unlike full-taxation regime, welfare increases under zero-rated regime, but at the cost of relative modest increase in the tax revenue.

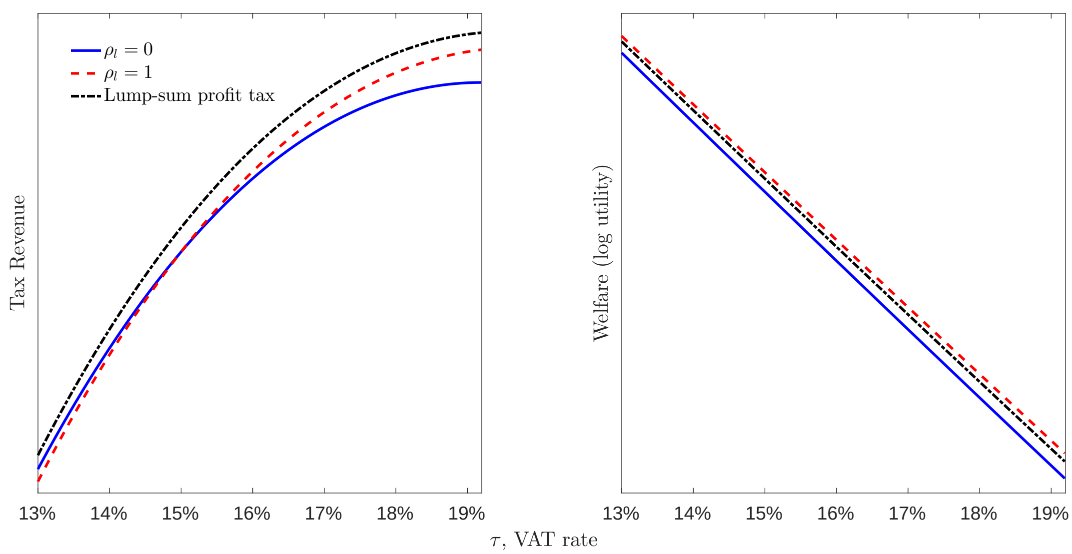

Figure 4: Change in Tax Revenue and Welfare with the VAT rate under Different Tax Regimes in Loan Services

We next change the VAT rate and examine the effects of the tax policy regime shift on total tax revenue and welfare. In figure 4, we compare tax revenue and welfare under the three tax regimes for loan services. First, welfare uniformly decreases with the VAT rate under each regime. For a given tax rate, the full-taxation without the profit tax yields higher utility than the full-taxation with the profit tax and both of them yield higher utility than the exempt treatment. However, the tax revenue does not uniformly change with the VAT rate and total tax revenue exhibits a Laffer curve under each regime. An interesting observation about the tax revenues is that while the full-taxation with profit tax generates more revenue than the other regimes, the exempt treatment regime generates more tax revenue than the full-taxation regime when the VAT rate is below 15%. In other words, repealing the exempt treatment comes at a cost which manifests itself on the total tax revenue.

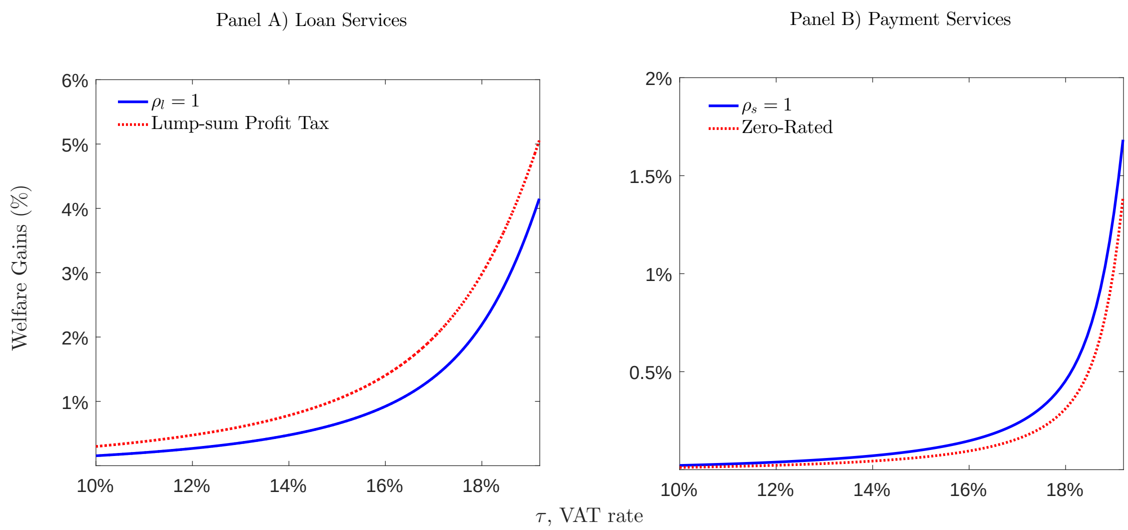

Figure 5: Tax Revenue Neutral Welfare Gains Relative to Exempt Treatment under Different Tax Regimes

To better evaluate the welfare gains under different tax regimes, we calculate welfare gains under each regime for a given total tax revenue. In particular, for a given tax rate, we first calculate total tax revenue under exempt treatment. Then, we change the VAT rate under the other regimes to keep the tax revenue constant. In Figure 5 Panel A, we report the percentage increase in welfare relative to exempt treatment. A tax-neutral analysis implies that welfare gains are larger under full-taxation. Compared to the exempt treatment regime, welfare increases with the tax rate and reaches to around 4% gain at $\tau = 0.192$, which is the peak of the Laffer curve.[15] A more striking result is that full-taxation with profit tax implies even larger welfare gains reaching to a maximum of 5%. This result suggests that the exempt treatment regime in loan services are welfare worsening

[15]We restrict the range of tax rates and focus on the tax rates that are to the left of the Laffer curve.

up to the point that increases the cost of financing for per unit of capital. However, larger fixed financing cost under exempt treatment is welfare improving, because it keeps low productive firms outside of the market. We achieve this outcome here by introducing a lump-sum profit tax. The implied welfare gain is one percentage point larger than the full-taxation without this lump-sum profit tax.

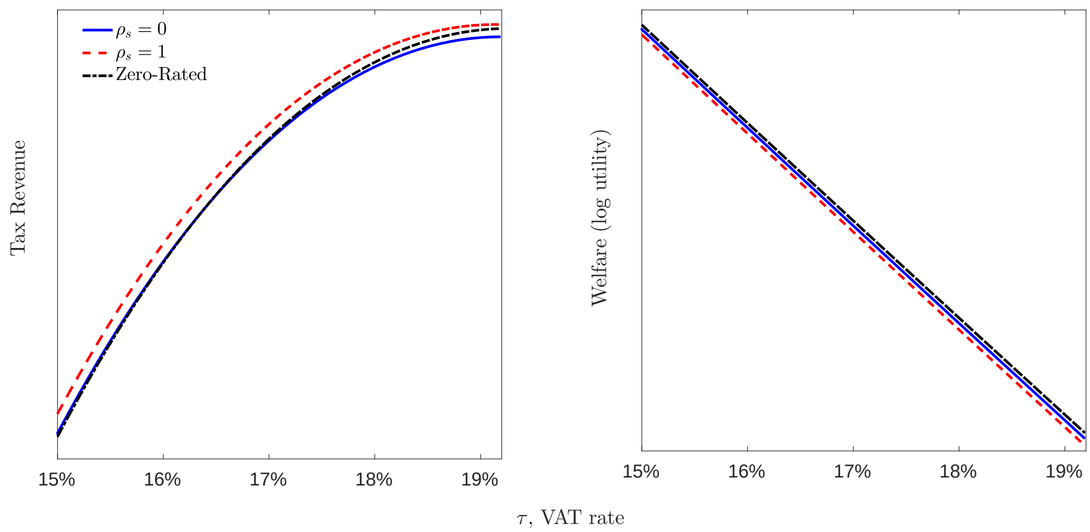

Figure 6: Change in Tax Revenue and Welfare with the VAT rate under Different Tax Regimes in Payment Services

Finally, we perform the same analysis above for different tax regimes in the payment services. Figure 6 presents our findings. While full-taxation reduces welfare, the zero-rated regime is welfare improving at every VAT rate. However, these welfare gains comes at a cost and associated with a decrease in total tax revenue at low VAT rates. In the case of full-taxation, despite welfare gains, total tax revenue increases significantly relative to the exempt treatment regime.

As in the case of loan services, we perform a tax-neutral policy analysis for payment services. Figure 5 Panel B presents our findings. Under zero-rated tax regime, welfare is higher compared to the exempt treatment case, where consumers are under-taxed, and increase with the VAT rate reaching to a 1.5% gain at around 19%. At first, this finding suggests that B2C transactions should not be taxed. However, welfare gains are even higher with $\rho_s = 1$ where the consumer is effectively taxed at rate τ. The reason is that payment services are part of total final output that use the resources of the economy. By taxing the consumers for their use of the payment services at the ongoing VAT rate, the government can afford to lower the VAT rate and redirect the resources to the production of the consumption good. Therefore, welfare gains are higher under this regime and reaches to a maximum of 2%.

6 Conclusion

The treatment of financial institutions under the VAT is a very complex issue in public finance. For various technical and practical reasons, financial services are generally exempt from VAT in most countries and this policy regime creates several distortions for the real economy. While some related literature has discussed the importance of these distortions, the precise transmission mechanisms of these distortions, and their net impact on the real economy and welfare have not been explored in a rich general equilibrium setting.

We achieve this goal in a dynamic general equilibrium framework with a banking sector. We show these mechanisms in model and evaluate many of the aforementioned distortions. The model accommodates both extensive and intensive margin and we find that the effects on the extensive margin are sizable and have significant implications for policy evaluations.

Using our calibrated model we performed policy analysis under different tax regimes. We find that repealing exempt treatment and moving toward full taxation in loan services lowers entry costs into the production sector, induces the entry of new firms, and lowers capital costs for all the firms. Overall, welfare is higher under full taxation in a tax-neutral regime switch. Welfare gains are even bigger when a lump-sum tax on profits is imposed so that low productive firms are kept outside of the market.

Regarding B2C transactions, we find that full-taxation reduces welfare for a given tax rate, but welfare improving in a tax-neutral tax policy regime switch. Zero-rating of payment services are also welfare improving over exempt treatment, but not as much as the full taxation regime. We conclude that taxing B2C transactions is welfare improving.

References

[1] Auerbach, A. J. and R. H. Gordon (2002). Taxation of financial services under a VAT. American Economic Review, 92(2), 411416.

[2] Bartelsman, E., S. Scarpetta, and F. Schivardi (2003). Comparative Analysis of Firm Demographics and Survival: Micro-Level Evidence for the OECD Countries. OECD Economics Department Working Papers 348, OECD Publishing.

[3] Bettendorf, L., M. P. Devereux, A. van der Horst, S. Loretz, and R. A. de Mooij (2009). Corporate tax harmonization in the eu. Working Papers 0932, Oxford University Centre for Business Taxation.

[4] Bird, R. M., P.-P. Gendron, and J. L. Rotman (2005). Vat revisited a new look at the value added tax in developing and transitional countries. USAID Report.

[5] Boadway, R. and M. Keen (2003). Theoretical Perspectives on the Taxation of Capital In- come and Financial Services, Chapter 2, pp. 3180. World Bank and Oxford University Press.

[6] Burke, A., H. Gorg, and A. Hanley (2008). The impact of foreign direct in- vestment on new firm survival in the UK: evidence for static versus dynamic industries. Small Business Economics 31(4), 395407.

[7] Buttner, T. and K. Erbe (2012). Revenue and welfare eects of financial sector vat exemption. TaxFACTs.

[8] Caminal, R. (2002). Taxation of banks: A theoretical framework. UFAE and IAE Working Papers 525.02, Unitat de Fonaments de lAnalisi Economica (UAB) and Institut dAnalisi Economica (CSIC).

[9] Chia, N.-C. and J. Whalley (1999). The tax treatment of financial intermediation. Journal of Money, Credit and Banking 31 (4), 70419.

[10] Chisari, O., A. Estache, and G. Nicodeme (2013). E ciency and equity eects of taxing the financial sector: Lessons from a cge model for belgium. Working Papers ECARES ecares 2013-01, ULB Universite Libre de Bruxelles.

[11] Denis, D. J. and V. T. Mihov (2003). The choice among bank debt, non-bank private debt, and public debt: evidence from new corporate borrowings. Journal of Financial Economics 70(1), 328.

[12] EU Commision (2010). Taxation papers: Financial sector taxation. Technical report, European Commission.

[13] Feenstra, R. C. (1994). New product varieties and the measurement of interna- tional prices. American Economic Review 84 (1), 15777.

[14] Gendron, P.-P. (2007). Value added tax treatment of financial services: An assessment and policy proposal for developing countries. International Tax Program Papers 0701, International Tax Program, Institute for International Business, Joseph L. Rotman School of Management, University of Toronto, http://ideas.repec.org/p/ttp/itpwps/0701.html.

[15] Greenwood, J., Z. Hercowitz, and G. W. Huffman (1988). Investment, capacity utilization, and the real business cycle. The American Economic Review 78 (3), pp. 402417.

[16] Grubert, H. and J. B. Mackie III (2000). Must financial services be taxed under a consumption tax? National Tax Journal 53 (1), 2340.

[17] Helmers, C. and M. Rogers (2008). Innovation and the Survival of New Firms Across British Regions. Economics Series Working Papers 416, University of Oxford, Department of Economics.

[18] Helpman, E., M. J. Melitz, and S. R. Yeaple (2004). Export versus fdi with heterogeneous firms. American Economic Review 94 (1), 300316.

[19] Houston, J. and C. James (1996). Bank information monopolies and the mix of private and public debt claims. Journal of Finance 51 (5), 186389.

[20] Huizinga, H. (2002). A European vat on financial services? Economic Policy 17(35), 497534.

[21] IMF (2010). A fair and substantial contribution by the financial sector. Report, IMF.

[22] Jack, W. (2000). The treatment of financial services under a broad-based consumption tax. National Tax Journal 53(4, Part 1), 841851.

[23] Johnson, S. A. (1997). An empirical analysis of the determinants of corporate debt ownership structure. Journal of Financial and Quantitative Analysis 32 (1), 4769.

[24] Kerrigan, A. (2010, March). The elusiveness of neutrality why is it so di cult to apply vat to financial services? International VAT Monitor 21 (2), 103112.

[25] Kleven, H. J., W. F. Richter, and P. B. Sorensen (2000). Optimal taxation with household production. Oxford Economic Papers 52, 584594.

[26] Lockwood, B. (2010). How should financial intermediation services be taxed? Working Paper 1014, Oxford University Centre for Business Taxation, http://ideas.repec.org/p/btx/wpaper/1014.html.

[27] McKenzie, K. J. and M. Firth (November). The GST and financial services: Pausing for perspective. Technical report, The GST at 20: The Future of Consumption Taxes in Canada, School of Public Policy, University of Calgary.

[28] Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. Econometrica 71(6), 16951725.

[29] OECD (2012). Entrepreneurship at a glance 2012. Technical report, OECD.

[30] Piggott, J. and J. Whalley (2001). Vat base broadening, self supply, and the informal sector. The American Economic Review 91 (4), 10841094.

[31] Poddar, S. (2003). Chapter 12 Consumption Taxes: The Role of the Value-Added-Taxes, Chapter 12, pp. 345380. World Bank and Oxford University Press.

[32] Russ, K. N. and D. Valderrama (2012). A theory of bank versus bond finance and intra-industry reallocation. Journal of Macroeconomics 34(3), 652 673.

[33] Schenk, A. S. (2009). Taxation of financial services (including insurance) under a united states value added tax. Technical Report 1520704, http://ssrn.com/paper=1520704.