CESifo WORKING PAPERS

6675
2017

September 2017

Indirect Taxation of Financial Services

Vidar Christiansen

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 6675*
Category 1: Public Finance

Indirect Taxation of Financial Services

Abstract

An important question is whether VAT exemption of financial services is a desirable property or whether it is justified only due to practical and administrative necessity. This paper singles out a number of financial services for discussion of this issue in a context allowing for other taxes and other preexisting distortions. It discusses taxation of intermediation that facilitates savings and borrowing, payment services and currency exchange. It also elaborates on the distortionary effects of taxing intermediate goods due to VAT exemption with focus on exports and consumer prices.

JEL-Codes: H200, H210, H220.

Keywords: financial services, indirect taxation, value added tax, VAT exemptions.

Vidar Christiansen
Department of Economics
University of Oslo
P.O. Box 1095 Blindern
Norway – 0317 Oslo
vidar.christiansen@econ.uio.no

This paper is part of the research at Oslo Fiscal Studies at the Department of Economics, University of Oslo, supported by the Research Council of Norway. I thank participants in the OFS Workshop on Indirect Taxation, the Tax Day Conference of the Max Planck Institute for Tax Law and Public Finance, the CESifo Area Conference on Public Sector Economics and the 73rd Annual Congress of IIPF in 2017, for helpful comments. In particular, I have benefitted from lengthy discussions with Stephen Smith.

1. Introduction

An unresolved issue in indirect taxation is how to tax financial services. The aftermath of the financial crisis has witnessed extensive interest in the taxation of the financial sector prompted in particular by concerns with excessive risk-taking and other conceivable weaknesses inherent in financial markets[1]. Besides this market failure approach, a central issue of interest has been the (missing) role of the value added tax and, more generally, indirect taxation of financial services. The exemption of many financial services from value added taxes has prompted discussions of whether financial services should be subject to VAT on the same basis as other goods and services, so long as satisfactory mechanisms can be designed to implement this. It is a widespread view in policy documents, public debate and parts of the academic literature that failure to tax financial services creates distortions on the grounds that financial services are being tax favoured compared to other goods and services. According to this view, our main task should be to search for feasible mechanisms for taxing financial services. Where a standard VAT (usually referred to as the credit-invoice method) is found impracticable one should search for an operable proxy. For instance, this seems to be the overriding argument in the report on the financial crisis from a government-appointed committee in Norway (Finanskriseutvalget; see NOU (2011: 1)[2]), and the Norwegian government's proposal to levy new taxes on the financial sector similar to those introduced in a number of other countries. For instance, a number of EU countries have successively made use of the optional clause in the VAT Directive to introduce taxes on various kinds of tax bases related to financial services. (See e.g. Ernst and Young , 2009, and Keen et al., 2010).

The argument is simple. When most commodities are taxed exempting some will create distortions by inducing substitution from taxed to untaxed goods and services. A caveat is that when the tax exempt sector purchases inputs on which a VAT is charged this VAT will be embedded in the sales price of the commodities supplied by the sector. When the tax exempt sector itself delivers an intermediate good a VAT is added to the price in later stages of the production and trade chain. Then any tax on the purchased inputs will be part of the tax base in later stages and VAT will be imposed on VAT – known as cascading. The prevailing view is therefore that financial services are tax favoured when delivered to consumers and over-taxed when delivered to businesses, in turn implying over-taxation of the final commodities in which these intermediate goods are embedded. A further distortion is the incentive for tax exempt businesses to choose

[1] For surveys of relevant issue, see e.g. Keen (2011) and Devereux (2011).
[2] Other Norwegian references are Meld. St. 4 (2015-2016) and NOU (2014: 13)).

in-house production rather than out-sourcing even when the former is more costly since a tax saving is obtained.

The academic literature conveys mixed conclusions regarding taxation of financial services. Auerbach and Gordon (2002) argue that "in principle, the VAT should apply to resources devoted to financial transactions in the same way as it does in other sectors". Grubert and Mackie (2000) question whether an explicit tax on financial services is justified, and, referring to a variety of financial services, they write: "We find that if the government fully taxes all direct consumption of goods and services, then also taxing these financial services will distort the consumer's decisions." A recurring argument in their discussion is that financial services are not consumption goods that directly yield the consumer utility. Boadway and Keen (2003) dismiss this argument saying that many items are commonly being taxed, or recommended to be so, even though they are clearly not final objects of utility. A discussion about what a consumption good is may easily become either semantic or philosophical. A more appropriate approach than just labelling items to determine tax liability or exemption is to consider all goods specified in a model as potential tax objects. Deriving optimal taxes will then determine the desirability of taxing the various items.

This paper will focus on VAT issues being concerned with potential tax distortions rather than the conceivably tax-correctable market failures highlighted in discussions of the financial crisis. There is a wide range of financial services (financial intermediation, payment services, currency trade, etc.), and key questions are whether they should all be subjected to "standard taxation" according to theoretical principles or whether there are arguments for taxing some or all categories of financial services differently from the general rate of VAT. Where, on theoretical grounds, no VAT is needed for efficiency one can in fact call off the search for proxies prompted by the view that in principle the VAT should be ubiquitous.

The overall view conveyed by the paper, is that there is limited scope for drawing general conclusions about how to tax a general category of items called 'financial services'. There is a need for considering the special properties of each financial service market. We shall discuss indirect taxation of a variety of major financial services. An exception is insurance, which is left out because it is a big topic with many aspects that should rather be dealt with separately. The first topic is taxation of financial intermediation facilitating savings and borrowing. We reconsider the taxation of fixed fees and margin-based fees, and argue that we should consider these taxes in a context that allows for direct taxes on interest as both kinds of taxes have similar effects. In Section 3 we take a closer look at the taxation of intermediate goods. Section 4 addresses taxation of payment services when some services are charged for explicitly while the costs of others are hidden as part of the general commodity price. Section 5 is devoted to a discussion of currency trade related to imports

of goods. Finally, a brief paragraph highlights distortions between in-house production and out-sourcing. Section 7 concludes.

2. Financial intermediation

2.1 A benchmark model.

An important category of financial services is financial intermediation facilitating savings and borrowing. To address main features of these financial services, we take as our point of departure a framework which is close to a simple model in Boadway and Keen (2003). To establish a benchmark, we assume that the representative consumer earns an exogenous income Y_j and pays a lump-sum tax T_j in period j (=1,2). First period income can be used for consumption C_1 and (positive or negative) savings, S. In the second period the exogenous income plus interest and first period savings are used for consumption C_2. Also assume that the consumer incurs a fixed cost K in period 1 and Q in period 2 charged by a financial institution (bank) when the consumer establishes a bank account or takes a loan. (Assuming the same costs in either case is a pure simplification.) There is a fixed return to capital, say in the world market, and a fixed transaction cost in the banking sector per unit of savings or borrowing. Denote by ρ_i the interest rate inclusive of transaction costs in the case of borrowing (i=B) and net of transaction costs in the case of positive savings (i=S); $\rho_B > \rho_s$. The consumer has preferences over consumption bundles given by a utility function $U(C_1, C_2)$.

The economy is described by the following set of equations

$$Y_1 - T_1 = C_1 + K + S \tag{1}$$

$$S = Y_1 - T_1 - C_1 - K \tag{1'}$$

$$S(1+\rho_i) + Y_2 - T_2 = C_2 + Q \tag{2}$$

$$(Y_1 - T_1 - K)(1+\rho_i) + Y_2 - T_2 - Q = C_1(1+\rho_i) + C_2 \tag{3}$$

$$U(C_1, C_2) \tag{4}$$

$$\frac{U_1}{U_2} = 1 + \rho_i \tag{5}$$

where, for notational simplicity, $U_i = \partial U / \partial C_i$ (i=1,2). (1) – (2) are budget (resource) constraints for the respective periods, and (3) is the intertemporal budget constraint. (5) is the optimality condition for the first best allocation equating the marginal rate of substitution to the marginal rate of transformation. Optimal consumption levels, savings, and utility level are determined by equations (2) – (5).

Now introduce a tax with rate τ on consumption. To relate our analysis to the discussion of VAT exemptions, we shall refer to this tax as VAT even though the characteristics defining a VAT do not appear in our model since, for simplicity, it neglects the input-output structure. We shall also consider the option to have a similar tax with rate $\delta\tau$ on the fixed costs K and Q where δ is one or zero. These indirect taxes replace the lump-sum taxes.

Now the economy is described by the following set of equations.

$$Y_1 = (1+\tau)C_1 + (1+\delta\tau)K + S \tag{6}$$

$$S = Y_1 - (1+\tau)C_1 - (1+\delta\tau)K \tag{6'}$$

$$S(1+\rho_i) + Y_2 = (1+\tau)C_2 + (1+\delta\tau)Q \tag{7}$$

$$\left(Y_1 - (1+\tau)C_1 - (1+\delta\tau)K\right)(1+\rho_i) + Y_2 = (1+\tau)C_2 + (1+\delta\tau)Q \tag{8}$$

Rearranging terms this is equivalent to

$$\left[(Y_1 - (1+\delta\tau)K)(1+\rho_i) + Y_2 - (1+\delta\tau)Q\right] / (1+\tau) = (1+\rho_i)C_1 + C_2 \tag{9}$$

$$\frac{U_1}{\mathrm{U}_2} = 1+\rho_i \tag{10}$$

The consumption levels are then determined by (9) and (10). We note that when taxes are chosen such that the left hand side of (9) and the left hand side of (3) are equal the representative consumer will choose the same consumption levels in the two tax regimes. In particular, there is no distortion at the margin, i.e., (10) holds in either case. Secondly, we can conclude that these results hold for any tax on the costs *K* and *Q* as long as they are indeed fixed. Boadway and Keen (2003) consider a model which is a special case of the one above (in their model positive savings are assumed and $Y_2 = Q = 0$), and conclude: "Financial services charged for as a fixed fee should be taxed". This conclusion is referred to as "a simple benchmark for policy". While a tax on the fixed costs is harmless from a social efficiency perspective in the model above, there seems to be no reason why the tax would indeed be strictly beneficial. Where a direct lump-sum tax is available (as at least one element of the tax system) there is no reason also to impose a tax on the fees charged by financial institutions for fixed costs. However, whether these costs are indeed incurred may be endogenous. In Boadway and Keen *K* is a cost of acquiring a savings account allowing the consumer to save up for the next period. This means that the benefits from savings must be large enough to outweigh this cost. Where this is the case with a small margin a tax on *K* may reverse the decision to save, and there is a tax distortion of savings behaviour. A tax on *K* is no longer a lump sum tax. A further concern may be that a tax on the "fixed" fee may induce substitution from fixed fee to a margin-based fee.

2.2 Taxation of margin-based fees

After having discussed the taxation of *K* and *Q*, let us now disregard these cost items. Unlike Boadway and Keen that assumed positive savings, we choose to consider the case where the consumer borrows. Let

borrowing be denoted by *B*. We shall now take into account that there is an income tax on exogenous income and interest, where interest on debt is deductible[3]. For simplicity, we assume there is a single tax rate denoted by *t*. The interest paid on debt is a world market interest rate plus a spread, which is a margin based fee imposed to cover the cost of administering the loan. We write this as $\rho_B = i + k$ where *i* is the world market return to capital and *k* is the cost of the bank per unit of borrowing. We can interpret this as the value added of the bank. Now assume that a tax at rate $\delta\tau$ is imposed on *k*, where δ is one or zero. Then $\rho_B = i + (1+\delta\tau)k$.

Now the economy is described by the following set of equations

$$(1-t)Y_1 + B = (1+\tau)C_1 \tag{11}$$

$$B = (1+\tau)C_1 - (1-t)Y_1 \tag{11'}$$

$$(1-t)Y_2 - \mathrm{B}(1+\rho_B(1-\mathrm{t})) = (1+\tau)C_2 \tag{12}$$

$$(1-t)Y_2 - (1+\rho_B(1-\mathrm{t}))\left((1+\tau)C_1 - (1-t)Y_1\right) = (1+\tau)C_2 \tag{13}$$

Rearranging terms this is equivalent to

$$(1-t)Y_2 - (1+\rho_B(1-\mathrm{t}))(1-t)Y_1 = (1+\tau)C_2 + (1+\rho_B(1-\mathrm{t}))(1+\tau)C_1 \tag{14}$$

Utility maximisation subject to the budget constraint yields

$$\frac{U_1}{\mathrm{U}_2} = 1+\rho_B(1-t) = 1+(\mathrm{i}+(1+\delta\tau)\,\mathrm{k})(1-t) = 1+i+k-t(i+k)+\delta\tau k(1-\mathrm{t}) \tag{15}$$

When there is no VAT on the margin-based fee (the spread) there is clearly a distortion as

$$\frac{U_1}{\mathrm{U}_2} = 1+i+k-t(i+k) < 1+i+k$$

Borrowing is subsidised because of the deductibility of interest payments from income. *The effect of imposing a VAT (changing δ from zero to one) is then to alleviate the distortion as the VAT is counteracting the effect of the income tax.* The net tax wedge becomes $-t(i+k)+\delta\tau k(1-\mathrm{t})$. We note that the social optimality conditions are satisfied if $-t(i+k)+\delta\tau k(1-\mathrm{t}) = 0.$

If we assumed that the agent had positive savings, the interest rate net of the margin-based fee would be $\rho_S = i - k$. Then the VAT on the margin-based fee would tax and distort savings downwards. This insight

[3] Strictly speaking, there is no case for distortionary taxes when exogenous incomes can be taxed. Our model should be seen as a simplification of a model where incomes are endogenous earnings, allowing exclusive focus on taxation of savings and financial services.

prompted Boadway and Keen (2003) to conclude that the spread charge should be left untaxed, but also to point out that "the results need to be considered with some care." We may note that in the presence of a tax on the return to savings, the effect would even be to exacerbate the distortionary effect of the interest tax. However, it is important to recognise that the effect is opposite in the case of borrowing[4].

2.3 VAT and capital income taxation.

Let us consider in further depth the case where a potential tax on intermediation coexists with a capital income tax. We now turn to an optimal tax approach where income is treated as endogenous. We set up the following model, which is a version of the Corlett-Hague model (Corlett and Hague, 1953). We consider a representative agent in two periods, indexed 1 and 2, respectively. The agent consumes in both periods and works in one or both periods. *S* is savings in the first period where $S \geq 0$ or $S < 0$. In either case the consumer will be charged with the administrative cost of saving or borrowing (cost of intermediation services).

The agent is assumed to supply an amount *h* of labour which is remunerated by a wage rate w_1 in period 1 and by w_2 in period 1. We may assume that there is a fixed division of labour between the periods, including the special cases that labour is supplied only in one of the periods (i.e. $w_1 = 0$ or $w_2 = 0$). The utility function of the agent is $u(C_1, C_2, h)$. As above, there is a VAT rate τ and a tax rate *t* on interest (with interest on debt being deductible). Denote the interest rate by r. The budget constraints are $w_1 h = (1+\tau)C_1 + S$ and $(1+\tau)C_2 = w_2 h + (1+r-tr)S$, and the intertemporal budget constraint is then $(1+\tau)C_2 = w_2 h + (1+r-tr)\left(w_1 h - (1+\tau)C_1\right)$. It follows that the indirect utility function can be written $\mathrm{V}\left((1+r-tr)(1+\tau), (1+\tau), (1+r-tr)w_1 + w_2\right)$ where the three arguments can be interpreted as the price of C_1, the price of C_2 and the total wage rate assigned to *h*, respectively .

We let $r = i - \alpha k(1+\beta\tau)$, where

$\alpha = 1$ if the agent has positive savings

$\alpha = -1$ if the agent has negative savings

$\beta = 1$ if the financial sector is liable to pay VAT; otherwise $\beta = 0$

The agent will be charged for the administrative cost (including tax if any) through a larger interest if borrowing and a lower interest if lending.

[4] There are countries, e.g. Norway, where the representative household (households on aggregate), borrows more from banks than it lends to the banks.

Assume that the government faces an interest rate $z = i - \alpha k$ and has a tax revenue requirement

$$\mathrm{R} = \tau C_1(1+z) + \left(tr + \alpha\beta\tau k\right)\left(w_1 h - (1+\tau)C_1\right) + \tau C_2 . \tag{16}$$

We start out by assuming that there is no VAT on the financial services, i.e. $\beta = 0$. The optimal tax problem can be written as max $\mathrm{V}(\cdot)$ subject to the tax revenue requirement. We can formulate the Lagrange function

$$\begin{aligned}\Lambda &= \mathrm{V}\left((1+r-tr)(1+\tau),(1+\tau),(1+r-tr)w_1 + w_2, I\right)\\ &+\mu\left((1+z)\tau C_1 + (tr+\alpha\beta\tau k)\left(w_1 h - (1+\tau)C_1\right) + \tau C_2 - R\right)\end{aligned} \tag{17}$$

The argument *I* is a (hypothetical) exogenous income which enters the utility function to allow us to distinguish income effects. Let $\lambda = \partial V / \partial I$.

Optimising wrt τ we get the first order condition

$$\begin{aligned}\frac{\partial\Lambda}{\partial\tau} &= (\mu-\lambda)\left((1+r(1-\mathrm{t}))C_1 + C_2\right)\\ &-\mu\left((1+r)\tau\frac{\partial C_1}{\partial I} + \tau\frac{\partial C_2}{\partial I} - tr(1+\tau)\frac{\partial C_1}{\partial I} + trw_1\frac{\partial h}{\partial I}\right)\cdot\left(C_1(1+r(1-\mathrm{t})) + C_2\right)\\ &+\mu\left(\left((1+r)\tau - tr(1+\tau)\right)s_{11}\left(1+r(1-t)\right)\right) + \mu\tau s_{21}\left(1+r(1-t)\right)\\ &+\mu\left(\left((1+r)\tau - tr(1+\tau)\right)s_{12}\right) + \mu\tau s_{22} + trw_1 s_{01}\left(1+r(1-t)\right)\mu + trw_1 s_{02}\mu = 0\end{aligned} \tag{18}$$

where $\frac{\partial C_i}{\partial I}$ is an income derivative and s_{ij} are compensated (Slutsky) derivatives where i=0,1,2 and j=0,1,2 , where 0 refers to labour, and 1 and 2 to C_1 and C_2, respectively. Also differentiating wrt *t*, and defining $m = (1+r)\tau\frac{\partial C_1}{\partial I} + \tau\frac{\partial C_2}{\partial I} - tr(1+\tau)\frac{\partial C_1}{\partial I} + trw_1\frac{\partial h}{\partial I}$, we can rewrite the first order condition as

$$\begin{aligned}\frac{\partial\Lambda}{\partial t} &= \left(\mu-\lambda\right)rS\mu - mrS\mu + \left((1+r)\tau - tr(1+\tau)\right)s_{11}\left(-r(1+\tau)\right)\mu + \tau s_{21}\left(-r(1+\tau)\right)\mu\\ &+trs_{01}w_1(-r(1+\tau))\mu + \left((1+r)\tau - tr(1+\tau)\right)s_{10}(-rw_1)\mu + \tau s_{20}(-rw_1)\mu + trw_1 s_{00}(-rw_1)\mu = 0\end{aligned}$$

or equivalently

$$\begin{aligned}\frac{\partial\Lambda}{\partial t} &= \left(1-\frac{\lambda}{\mu}\right) - m + \frac{\left((1+r)\tau - tr(1+\tau)\right)s_{11}\left(-r(1+\tau)\right)}{rS} + \frac{\tau s_{21}\left(-r(1+\tau)\right)}{rS} + \frac{trw_1 s_{01}w_1(-r(1+\tau))}{rS}\\ &+\frac{\left((1+r)\tau - tr(1+\tau)\right)s_{10}(-rw_1)}{rS} + \frac{\tau s_{20}(-rw_1)}{rS} + \frac{trw_1 s_{00}(-rw_1)}{rS} = 0\end{aligned} \tag{19}$$

As is well known from the Corlett-Hague model, the optimum will trade off the distortion of the consumption bundle (C_1, C_2) and the distortion of labour supply. But our main concern in the current paper is how to tax the financial services rather than the characteristics of the optimum above. Taking VAT exemption in the financial sector $(\beta = 0)$ as our point of departure, our next step is to consider the effect of introducing a VAT on *k*. From above we see that

$$\frac{dr}{d\beta} = -\alpha k \tau$$

From the envelope theorem

$$\begin{aligned}
&\frac{\partial}{\partial \beta}\Lambda = -\lambda S(1-\mathrm{t})\alpha\,\mathrm{k}\,\tau + \mu(-\mathrm{t}\,\alpha\,\mathrm{k}\,\tau + \alpha k\tau)\mathrm{S} - \mu\,\mathrm{m}(1-\mathrm{t})\alpha S\,\mathrm{k}\,\tau \\
&+\mu((1+r)\tau + tr(1+\tau))\,\mathrm{s}_{11}(1-\mathrm{t})(1+\tau)(-\alpha\,\mathrm{k}\,\tau) + \mu((1+r)\tau + tr(1+\tau))\,\mathrm{s}_{10}(1-\mathrm{t})\,\mathrm{w}_1(-\alpha\,\mathrm{k}\,\tau) \\
&+\mu tr w_1\,\mathrm{s}_{01}(1+\tau)(1-\mathrm{t})(-\alpha\,\mathrm{k}\,\tau) + \mu tr w_1\,\mathrm{s}_{00}(1-\mathrm{t})\,\mathrm{w}_1(-\alpha\,\mathrm{k}\,\tau) + \mu\tau s_{21}(1-\mathrm{t})(1+\tau)(-\alpha\,\mathrm{k}\,\tau) \\
&+\mu\tau\,\mathrm{s}_{20}(1-\mathrm{t})\,\mathrm{w}_1(-\alpha\,\mathrm{k}\,\tau)
\end{aligned}$$

or equivalently

$$\begin{aligned}
&\frac{\partial}{\partial \beta}\Lambda = S(1-\mathrm{t})\alpha\,\mathrm{k}\,\tau\mu[1-\lambda/\mu-\mathrm{m} \\
&+\frac{((1+r)\tau - tr(1+\tau))\,\mathrm{s}_{11}(1-\mathrm{t})(1+\tau)(-\alpha\,\mathrm{k}\,\tau)}{S(1-\mathrm{t})\alpha\,\mathrm{k}\,\tau} + \frac{(1+r)\tau\,\mathrm{s}_{10}(1-\mathrm{t})\,\mathrm{w}_1(-\alpha\,\mathrm{k}\,\tau)}{S(1-\mathrm{t})\alpha\,\mathrm{k}\,\tau} - \frac{tr(1+\tau))\,\mathrm{s}_{10}(1-\mathrm{t})\,\mathrm{w}_1(-\alpha\,\mathrm{k}\,\tau)}{S(1-\mathrm{t})\alpha\,\mathrm{k}\,\tau} \\
&+\frac{tr w_1\,\mathrm{s}_{01}(1+\tau)(1-\mathrm{t})(-\alpha\,\mathrm{k}\,\tau)}{S(1-\mathrm{t})\alpha\,\mathrm{k}\,\tau} + \frac{tr w_1\,\mathrm{s}_{00}(1-\mathrm{t})\,\mathrm{w}_1(-\alpha\,\mathrm{k}\,\tau)}{S(1-\mathrm{t})\alpha\,\mathrm{k}\,\tau} + \frac{\tau s_{21}(1-\mathrm{t})(1+\tau)(-\alpha\,\mathrm{k}\,\tau)}{S(1-\mathrm{t})\alpha\,\mathrm{k}\,\tau} + \frac{\tau\,\mathrm{s}_{20}(1-\mathrm{t})\,\mathrm{w}_1(-\alpha\,\mathrm{k}\,\tau)}{S(1-\mathrm{t})\alpha\,\mathrm{k}\,\tau}]
\end{aligned} \tag{20}$$

We find that when $\frac{\partial \Lambda}{\partial t} = 0$, then $\frac{\partial}{\partial \beta}\Lambda = 0$.

When there is an optimal tax on interest imposing VAT on the financial service makes no difference.

Above we have assumed that the tax on interest is set based on a simple intertemporal model. In practice, there may be other concerns involved. If it is considered desirable to discourage borrowing this could be done by lowering *t* to diminish the effect of interest deductibility. If this is difficult due to other concerns, there may be a case for abolishing the VAT exemption. A case may be where the rate of the capital income tax *t* is linked to the tax rate on profits (as in Norway) so that a lower *t*, unlike a tax on the intermediation services, will imply a more lenient taxation of profits.

3. Missing refund of VAT on inputs

3.1 Taxation of intermediate goods

Most commodities, including financial services, are not only domestic services sold to the households but are also used as intermediate goods in the production of other goods. For instance, businesses will use financial services as inputs in their activities. In order to examine the economy-wide effects of having a VAT exempt sector, we need to consider both its sales to consumers and its interaction with other sectors. In doing so, we should allow for the fact that there are sectors producing non-tradables as well as importing and exporting sectors. To capture all these features we shall establish a model with four sectors numbered 1, 2, 3, 4. This is sufficient to illustrate the effects of interest. In general, each sector purchases inputs from the other sectors besides using some primary factor, say labour, which generates a value added in the sector. The input-output structure is described by input-output coefficients a_{ij} which is the amount of the sector *i* good used as input per unit produced in sector j. Denote by v_i the value added and by p_i the (producer) price per unit produced in sector *i* (i=1,2,3,4). We shall assume that good 1 is imported at a fixed world market price. Commodity 2 and commodity 3 are domestically produced non-tradables. Commodity 4 is exported at a fixed world market price. p_1 and p_4 are then treated as exogenous while p_2 and p_3 are endogenous in the model. Sector 2 is assumed to be VAT exempt while all other sectors are VAT registered. Denote by τ the VAT rate.

Producer prices are then given by:

p_1 is exogenous

$$\mathrm{p}_2 = v_2 + (1+\tau)p_1a_{12} + (1+\tau)p_3a_{32} + (1+\tau)p_4a_{42} \tag{21}$$

$$\mathrm{p}_3 = v_3 + p_1a_{13} + p_2a_{23} + p_4a_{43} \tag{22}$$

p_4 is exogenous

In the non-tradable sectors the price is assumed to be reflect the value added and the cost of intermediate inputs from other sectors. While the VAT on inputs is refunded in VAT registered sectors it is not in the VAT exempt sector 2. We assume that there is fixed value added per unit in sector 2 and sector 3. In order to endogenise exports, we assume that sector 4 has a production technology such that the value added required to produce x_4 units is $v_4(x_4)$ where $v_4'(x_4) > 0$. The marginal cost in sector 4 is then $v_4'(x_4) + p_1a_{14} + p_2a_{24} + p_3a_{34}$, which at the supplier's optimum is equated to the price:

$$p_4 = v_4'(x_4) + p_1a_{14} + p_2a_{24} + p_3a_{34} \tag{23}$$

Rearranging terms, we can write (21) and (22) as

$$\mathrm{p}_2 - (1+\tau)p_3a_{32} = v_2 + (1+\tau)p_1a_{12} + (1+\tau)p_4a_{42}$$

$$-p_2 a_{23} + \mathrm{p}_3 = v_3 + p_1 a_{13} + p_4 a_{43}$$

Solving for the endogenous prices we get

$$p_2 = \frac{v_2 + (1+\tau) p_1 a_{12} + (1+\tau) p_4 a_{42} + (1+\tau) a_{32} \left(v_3 + p_1 a_{13} + p_4 a_{43} \right)}{1-(1+\tau) a_{32} a_{23}} \tag{24}$$

$$p_3 = \frac{v_3 + p_1 a_{13} + p_4 a_{43} + a_{23} \left(v_2 + (1+\tau) p_1 a_{12} + (1+\tau) p_4 a_{42} \right)}{1-(1+\tau) a_{32} a_{23}} \tag{25}$$

Recalling $p_4 = v_4^{'}(x_4) + p_1 a_{14} + p_2 a_{24} + p_3 a_{34}$, we note that where, due to VAT exemption, non-refunded taxes are passed through to the input prices paid by the exporting sector its marginal cost is distorted upwards. There is in fact imperfect zero rating of exports. The private cost in sector 4 exceeds the social cost when the sector is dependent on deliveries from the tax exempt sector. There is a downward distortion of exports. An increase in τ will lower output and exports from the sector and further exacerbate this distortion[5].

Our next question is whether a VAT system with exemptions distorts consumer prices. Denote by q_i the consumer price of commodity *i*. Considering relative consumer prices, we find

$$\frac{q_2}{q_1} = \frac{p_2}{(1+\tau) p_1} = \frac{v_2 / (1+\tau) + p_1 a_{12} + p_4 a_{42} + a_{32} \left(v_3 + p_1 a_{13} + p_4 a_{43} \right)}{\left(1-(1+\tau) a_{32} a_{23} \right) p_1} \tag{26}$$

The exempt sector 2 does not charge VAT while the non-exempt sector 1 does. We note that there are two opposing distortionary effects on consumer prices. Goods from the exempt sector are under-priced when no VAT is charged on the value added embedded in goods sold to consumers. However, where the VAT registered sector 3 charges VAT on intermediate goods delivered to sector 2, which in turn delivers inputs back to sector 3 and so forth ($a_{32} a_{23} > 0$) there is pass-through of non-refunded VAT on inputs, the partial effect of which is to raise the relative price of commodity 2. One will normally expect the former effect to dominate.

$$\frac{q_3}{q_1} = \frac{(1+\tau) p_3}{(1+\tau) p_1} = \frac{v_3 + p_1 a_{13} + p_4 a_{43} + a_{23} \left(v_2 + (1+\tau) p_1 a_{12} + (1+\tau) p_4 a_{42} \right)}{\left(1-(1+\tau) a_{32} a_{23} \right) p_1} \tag{27}$$

We see that where there is pass-through of non-refunded VAT on inputs there is an upward price distortion.

It is often said that VAT exemption will imply too low prices to consumers and too high prices to businesses. However, prices to businesses are uninteresting as such. What matters for welfare are final consumer prices. We see that both consumer price ratios, determined in the domestic market, are

[5] Formally, this can be shown by simple comparative statics differentiating the condition above with respect to τ.

distorted due to the VAT exemption in sector 2. While there are to some extent opposing effects on the relative consumer price in the exempt sector, the relative consumer price in the non-exempt sector 3 is unambiguously biased upwards. How large this distortion is depends on the input-output structure[6]. It is smaller the smaller are the coefficient products $a_{23}a_{12}, a_{23}a_{42}$ and $a_{23}a_{32}$. Where the size is non-negligible, how serious the distortion is does of course depend on the substitution elasticity of demand[7].

$$\frac{q_4}{q_1} = \frac{(1+\tau)p_4}{(1+\tau)p_1} = \frac{p_4}{p_1} \tag{28}$$

Where pre-tax prices are exogenous world market prices there are no VAT distortions.

3.2 A proxy for VAT

Taking for granted that a VAT with refund of VAT on inputs is not feasible, the VAT exemption of certain services has motivated a search for proxies to make up for the missing VAT. There have been proposals and actual policy reforms to impose excises or to levy taxes on wages and profits in otherwise VAT exempt sectors. However, there is no refund of taxes on inputs as in fully-fledged VAT system. Two methods of this kind are known in the literature as the subtraction method and the addition method. In the financial sector, such a tax is often referred to as a FAT (financial activities tax).

Let us now examine the consequences of imposing a tax with these characteristics. We consider introducing a tax on the value added v_2 in the exempt sector. For convenience, we set the tax rate equal to the conventional VAT rate τ. We do this within the model considered above. Then v_2 is replaced by $(1+\tau)v_2$ in our formulas. Prices are then given by

p_1 is exogenous

$$\mathrm{p}_2 = (1+\tau)v_2 + (1+\tau)p_1a_{12} + (1+\tau)p_3a_{32} + (1+\tau)p_4a_{42}$$

$$\mathrm{p}_3 = v_3 + p_1a_{13} + p_2a_{23} + p_4a_{43}$$

p_4 is exogenous

$$p_2 = \frac{(1+\tau)v_2 + (1+\tau)p_1a_{12} + (1+\tau)p_4a_{42} + (1+\tau)a_{32}\left(v_3 + p_1a_{13} + p_4a_{43}\right)}{1-(1+\tau)a_{32}a_{23}} \tag{29}$$

[6] A paper studying empirically the effects of the input-output structure is Buettner and Erbe (2014).
[7] It is implicitly assumed that conceivable consumer price effects on labour supply (Corlett-Hague effects) are of no concern in our context.

$$p_3 = \frac{v_3 + p_1 a_{13} + p_4 a_{43} + a_{23}\left((1+\tau)v_2 + (1+\tau)p_1 a_{12} + (1+\tau)p_4 a_{42}\right)}{1-(1+\tau)a_{32}a_{23}} \tag{30}$$

Recalling the marginal cost of the exporting sector, $v_4'(x_4) + p_1 a_{14} + p_2 a_{24} + p_3 a_{34}$, we note that the proxy VAT will further raise the marginal cost of exports and exacerbate the downward export distortion.

$$\frac{q_2}{q_1} = \frac{p_2}{(1+\tau)p_1} = \frac{v_2 + p_1 a_{12} + p_4 a_{42} + a_{32}\left(v_3 + p_1 a_{13} + p_4 a_{43}\right)}{\left(1-(1+\tau)a_{32}a_{23}\right)p_1} \tag{31}$$

$$\frac{q_3}{q_1} = \frac{(1+\tau)p_3}{(1+\tau)p_1} = \frac{v_3 + p_1 a_{13} + p_4 a_{43} + a_{23}\left((1+\tau)v_2 + (1+\tau)p_1 a_{12} + (1+\tau)p_4 a_{42}\right)}{\left(1-(1+\tau)a_{32}a_{23}\right)p_1} \tag{32}$$

We see that it makes a difference if there is a proxy for the VAT in sector 2. Provided that there are interactions between sector 2 and the other non-tradable sector (sector 3) the relative consumer price in sector 2 is now unambiguously distorted upwards. But the distortion is minor if $a_{32}a_{23}$ is small. In that case, the proxy comes close to eliminating the previous downward distortion. Also the relative consumer price in sector 3 is upward distorted and even more than in the case of no proxy. We can conclude that in general relative prices of tradables are distorted upwards.

4 Taxing payment services

One of the tasks of the financial sector is to provide and facilitate payment services by making cards, cash and other modes of payment available and by processing transactions. Payments involve transaction costs due to use of resources by retailers, customers, financial institutions and possibly the government and central banks. The payment costs incurred by the various agents can affect the economic allocation in various ways. i. They increase the cost of market consumption in general which can affect the choice between market consumption and leisure or home production. ii. They can affect the number of transactions being made. iii. They are likely to affect the choice of modes of payment – use of cash, various cards, giro transactions, cheques, etc. iv. They can indirectly affect the choice of consumption bundle since payment costs and the preferred mode of payment varies with the size and possibly other characteristics of the purchase. For instance, typically cash is not used for large purchases. In the sequel, we shall focus on the choice of form of payment as it seems plausible that this is the major allocative effect of payment costs.

Some payment services requires use of time withdrawing cash, counting money, queuing in front of the cashier, etc. This raises the issue of efficient use of time. Lockwood and Yerushalmi (2017) address this issue considering the choice between cash and bank deposits where cash requires a time input. They show that there is a case for a tax on bank deposits reflecting the fact that deposits, by economising on household time, are complementary with leisure – a result in the vein of Corlett and Hague (1953).

Carrying out payments may also create externalities as it does when customers choose a slow form of payment increasing the waiting time of other customers queuing in front of the cashier, or inducing the shopkeeper to employ more cashiers. This phenomenon was presumably observed more frequently in the "old days" when people spent time on filling in and signing cheques.

Which modes of payment are available to the customers are, within certain legal constraints, determined by the vendors. For instance, a retailer may accept payment by card or not, or certain cards but not others, or may be willing to send an invoice or not. In some cases, there are economies of scale. If enough customers would like to use cards it may be worthwhile acquiring payment terminals. To a large extent, customers are free to choose payment medium, and it is this choice that is at the centre of interest in the following.

Customers can be charged directly or indirectly for payment costs. They are charged directly when they pay specifically for the payment cost. They are charged indirectly when the payment cost due to the activities of the seller or a financial institution is embedded in the commodity prices paid by all customers through a general mark-up. Ideally, each agent should face the transaction-specific cost he inflicts on society. In practice, this is not likely to happen even if new technology has taken us some steps towards this situation. An example is price discounts offered to those who choose a more efficient form of payment. But typically, the price is not being conditioned on the form of payment even if the transaction cost may differ. In the following, our focus is the choice between payment services paid for directly or indirectly. We shall abstract from time costs and externalities.

A financial institution will charge VAT or not depending on whether it is VAT registered or exempt.

The costs incurred by the shops and other VAT registered sellers, whether due to own activities or payments to financial institutions, will be subject to VAT irrespective of whether financial institutions are exempt or not.

Consider a transaction where a customer buys one unit. Assume there are two possible payment methods labelled *a* and *b*, respectively. The customer decides which one to use. Let the social cost of each one be C^a and C^b, respectively. If picking method *a* the customer has to pay a separate fee reflecting the cost of the payment service, i.e. he pays $p + C^a$ where *p* is the price net of the transaction cost C^a. When using method *b*, the payment cost is charged for through a mark-up of the price, and the customer pays *p* only. Some customers may still choose *a* if they find *a* more convenient. Suppose the convenience is equal to a benefit *B*, and suppose there is a distribution of people across values of *B* with density $f(B)$. Social efficiency would then imply that *a* should be chosen if $C^a - B < C^b$, i.e. net of benefit *a* is less costly than *b*. Private optimisation implies that *b* is chosen when $B < C^a$. The benefit is not large enough to offset the fee. This means that for values of *B* where $C^a - C^b < B < C^a$ method *b* is chosen despite being socially more costly than *a*. The benefit is large enough to make the net social cost of *a* smaller than that of *b* but not large enough to induce the customer to choose *a*. The choices of agents with these values of *B* will generate an excess burden, which for some value of *B* is $C^b - \left(C^a - B\right)$. The total excess burden can then be expressed as $\int_{C^a - C^b}^{C^a} \left[C^b - \left(C^a - B\right)\right] f(B) dB$. We may note that where $B < C^a - C^b < C^a$ method

b is the socially more efficient and will be chosen, and where $C^a - C^b < C^a < B$ it is method *a* that is the socially more efficient and will be chosen.

When there is a consumption tax (VAT) with a rate τ but where the fee for the payment service is tax exempt the customer is charged $(1+\tau)p + C^a$ if choosing *a* and $P = (1+\tau)p$ if deciding to use *b*. We note that the choice of the agent will remain the same as above. Then assume that a tax is imposed also on the financial service so that the fee-inclusive price of the item when choosing payment method *a* becomes $(1+\tau)p + (1+\tau)C^a$. Now *b* will be chosen when $B < (1+\tau)C^a$. The benefit is not large enough to compensate the agent for the tax-inclusive fee. Now the excess burden becomes $\int_{C^a - C^b}^{(1+\tau)C^a} \left[C^b - \left(C^a - B\right)\right] f(B)dB$. The distortion is exacerbated and the excess burden increases.

5. Taxing currency trade

Suppose a VAT-registered firm imports a good at a foreign price *p*. Let the exchange rate be *v* and suppose the cost of currency exchange is *kp* If the currency trader is tax exempt he charges *kp* and the import price including the cost of the financial service is *vp+kp*. Denote by *b* the cost (value added) of the activities carried out by the importer. Charging VAT when selling the good to a consumer, the importer will charge the tax-inclusive price $(1+\tau)(\mathrm{vp}+\mathrm{kp}+\mathrm{b})$. There is the same tax mark-up of the price as for other (domestically provided) goods.

If the currency trader charges VAT he pays $\tau\,\mathrm{kp}$ to the tax office and charges $(1+\tau)\,\mathrm{kp}$. The importer incurs a tax-inclusive cost $\mathrm{vp}+(1+\tau)\,\mathrm{kp}+\mathrm{b}$. $\tau\,\mathrm{kp}$ is refunded and the net cost faced by the importer is $\mathrm{vp}+\mathrm{kp}+\mathrm{b}$. The importer charges the tax-inclusive price $(1+\tau)(\mathrm{vp}+\mathrm{kp}+\mathrm{b})$. It makes no difference whether the currency trader is tax exempt or not. If the currency trader does not charge VAT, the importer will add VAT to the consumer price. If the currency trader charges VAT the importer will get a refund and add VAT to the consumer price. The only difference is who remits the VAT on the currency trade. This illustrates the more general insight that if there is a first stage of the production and distribution chain only delivering intermediate goods it makes no difference whether it is tax exempt or not.

A different case arises when currency exchange is a service purchased by consumers to be able to travel or do shopping abroad. Consumers will then directly import many goods and services from abroad on which no home country VAT is paid. Since VAT is levied on domestic goods and commercial import the missing VAT on cross border shopping will create a distortion. As a partial taxation of consumers' purchases abroad, VAT on currency exchange is therefore in principle efficiency enhancing but presumably with a minor quantitative impact.

6. Distortion of outsourcing versus in-house production.

Where a financial institution is VAT exempt it will have to pay VAT on intermediate goods and services that it acquires. When the institution is able to produce these goods and services by means of own employees the institution will have to compare the costs of in-house production versus outsourcing. When

VAT is imposed on purchased services but not on own production there is a distortion in favour of the latter. Whether this is a problem depends on the how efficient in-house production is compared to the efficiency of external suppliers. In-house production may be less efficient when external suppliers have a specialisation or a scale advantage. There is a non-negligible deadweight loss if in-house delivery is moderately less efficient. If it is slightly less efficient, the efficiency loss is minor. If it is considerably less efficient, it will never be chosen, and efficiency loss is not an issue.

7. Conclusion

This paper has singled out a number of financial services for further discussion. In particular, the paper has argued the need for addressing taxation of financial services in a context allowing for other taxes and conceivably other preexisting distortions.

Conclusions about how financial services should be taxed are mixed. Considering intermediation facilitating borrowing and saving, it has often been argued that fixed fees charged by the bank should be taxed while there is a weaker, if any, case for taxing the spread. These recommendations are based on the assumption that the latter is distortionary while the former is not. We have considered a number of qualifications. The former may not be entirely fixed as the customer may be free to choose whether to opt in or out of the savings or borrowing scheme. But assuming that a fixed fee for banking services is indeed equivalent to a lump sum tax, the two questions remain whether there is too little lump sum taxation in society and whether taxing bank fees is the efficient way to implement lump sum taxes. Moreover, that a tax is distortionary is not a sufficient argument against it as distortionary taxes are indeed accepted on a large scale. While a tax on the spread has typically been considered in isolation, the current paper argues that when a (direct) tax on interest is in place, already distorting savings and borrowing, a tax on the spread may exacerbate or alleviate pre-existing distortions. We have shown within a simple model that where an income tax on the return to capital and a consumption tax (VAT) exempting intermediation services are set optimally there is no gain from removing the exemption.

An important property of the VAT system is the refunding of VAT on intermediate goods purchased from other sectors. Exemptions imply that producers cannot reclaim the input VAT. Two implications have been highlighted. There will be a downward distortion of exports. And consumer prices in other sectors will be distorted. The extent to which there are distortions depend on the prevalence of world market prices versus domestically determined prices and the input-output structure of production.

When goods are traded payment services can be charged for explicitly or disguised as a mark-up of the price to which a VAT is added. Then there will be over-use of the payment services with hidden charges. A higher tax on the payment services bought separately will then exacerbate this distortion.

Whether the first stage of the production and distribution chain is VAT exempt makes no substantial difference for consumer prices at later stages but only determines who will remit the tax. An example is currency trade related to imports of goods. However, in principle it would be desirable to tax the consumers' purchases of currency exchange services.

Finally, as is well known, exemption from VAT will distort the choice between and out-contracting of tasks and tax-favoured in-house production.

References

Auerbach, A. J. and R. H. Gordon (2002). Taxation of Financial Services under a VAT. *The American Economic Review 92*, 411-416.

Boadway, R. and M. Keen (2003). Theoretical Perspectives on the Taxation of Capital Income and Financial Services: A Survey. In: Honohan, P. (Ed.), *Taxation of Financial Intermediation Theory and Practice for Emerging Economies.* New York: Oxford University Press

Buettner, T. and K. Erbe (2014). Revenue and welfare effects of financial sector VAT exemption. *International Tax and Public Finance 21*, 1028-1050.

Corlett, W. J. and D. C. Hague (1953). Complementarity and the excess burden of taxation. *Review of Economic Studies 21*, 21-30.

Devereux, M.P (2011). New Bank Taxes: Why and What Will Be the Effect? *European Tax Policy Forum*, URL: http://etpf.org/papers/64bank.pdf

Ernst and Young (2009). *Design and Impact of the "Option to Tax" System for application of VAT to financial services* . Report prepared for the European Banking Federation.

Grubert, H. and J. Mackie (2000). Must Financial Services be Taxed Under a Consumption. *Tax National Tax Journal 53*, 23-40.

Keen, M. (2011). Rethinking the Taxation of the Financial Sector. *CESifo Economic Studies 57*, 1-24.

Keen, M,, R. Krelove and J. Norregaard (2010). The financial activities tax. In S. Claessens, M. Keen and C. Pazarbasioglu (eds), *Financial sector taxation: the IMF's report to the G-20 and background material.* International Monetary Fund.

Lockwood, B. (2014). How should Financial Intermediation Services be Taxed? In: de Mooij, R. and G. Nicodème (Eds.), *Taxation and Regulation of the Financial Sector*. Massachusetts: The MIT press

Lockwood, B. and E. Yerushalmi (2017). How should transaction services be taxed? Paper presented at the 73rd Annual Congress of the International Institute of Public Finance.

Meld. St. 4 (2015-2016). Bedre skatt. En skattereform for omstilling og vekst. Oslo: Finansdepartementet

NOU (2014: 13). Kapitalbeskatning i en internasjonal økonomi. Oslo: Finansdepartementet

NOU (2011: 1). Bedre rustet mot finanskriser – Finanskriseutvalgets utredning. Oslo: Finansdepartementet