Working Paper

Decomposing the Margins of Transfer Pricing

Author(s):
Lassmann, Andrea; Zoller-Rydzek, Benedikt Marian Maximilian

Publication Date:
2019-01

Permanent Link:

https://doi.org/10.3929/ethz-b-000322881 →

ETH zürich

KOF Swiss Economic Institute

Decomposing the Margins of Transfer Pricing

Andrea Lassmann and Benedikt Zoller-Rydzek

KOF Working Papers, No. 450, January 2019

KOF

KOF

ETH Zurich
KOF Swiss Economic Institute
LEE G 116
Leonhardstrasse 21
8092 Zurich, Switzerland

Phone +41 44 632 42 39
Fax +41 44 632 12 18
www.kof.ethz.ch
kof@kof.ethz.ch

Decomposing the Margins of Transfer Pricing*

Andrea Lassmann† and Benedikt Zoller-Rydzek‡

February 4, 2019

Abstract

This paper examines the determinants and margins of profit shifting through transfer pricing. We develop a theory model, where transfer pricing patterns are governed by a generalized concealment cost function (CCF). Our empirical analysis draws on micro-level data about transaction-level imports, firm-level characteristics, as well as tax differentials between regions in Switzerland and countries abroad. We find, both theoretically and empirically, that more productive multinational firms deviate less from the arms' length price and trade lower quantities, compared to MNEs with lower productivity. Moreover, the decision of firms to engage in transfer pricing depends negatively on a fixed cost component in the CCF, as well as trade costs. The model allows us to estimate a theory-consistent concealment cost function, which can be used for counterfactual analysis.

JEL classification : F23; H25; H26; H32.

Key Words : Multinational firms; tax avoidance; tax havens; transfer pricing.

*We thank the Swiss Federal Customs Administration (EZV) for access to data on goods import transactions. We would like to thank Ron Davies, Dhammika Dharmapala, Peter Egger, Tim Schmidt-Eisenlohr, and audiences at the 3rd Annual MaTax Conference in Mannheim, the 9th FIW Research Conference, the Annual Conference of the ETSG in Helsinki, the Annual Congress of the International Institute of Public Finance in Lake Tahoe, the SNF Sinergia Workshop - Inequality and Globalization in Stoos, the University of Linz, and the University of Mainz for helpful comments.

†*Affiliation:* ETH Zurich. *Address:* ETH Zürich, KOF, Leonhardstrasse 21, LEE G 118, 8092 Zurich, Switzerland. E-mail: andrealassmann@gmail.com

‡*Affiliation:* ETH Zürich. *Address:* ETH Zürich, KOF, Leonhardstrasse 21, LEE G 131, 8092 Zurich, Switzerland. E-mail: rydzek@kof.ethz.ch.

1 Introduction

Tax harmonization and compliance initiatives have been prominent on the international institutional agenda in recent years. Several bilateral and multilateral tax-authority initiatives – in particular the OECD/G20 Base Erosion and Profit Shifting package (BEPS) – aim at collaborating on the matter of tax avoidance, which multinational enterprises (MNE) have been repeatedly accused of.

In this paper, we provide novel evidence on the forces that govern the patterns of transfer pricing, one of the main instruments for profit shifting. We define as the transfer price the price that MNEs set for intra-firm transactions of physical goods, compared to an arm's length price (ALP) for a transfer between independent parties. Intra-firm transfers create leeway for shifting profits for tax avoidance motives, and previous literature has indeed found evidence for transfer mispricing (see Bernard et al., 2006; Cristea and Nguyen, 2016; Davies et al., 2018; Liu et al., 2017). However, the factors that foster transfer mispricing have not been entirely disentangled. Specifically, research on (i) the question, which factors determine the decision of firms to engage in transfer pricing, i.e., the *extensive* margin, (ii) the extent of profits shifted, i.e., the *intensive* margin, and (iii) different transfer pricing strategies – firms may shift a given amount of profits to locations with lower corporate tax rates by either shipping small volumes at prices vastly different from the competitive ALP, by shipping large quantities at prices very similar to the ALP, or by any combination of the two – is scarce. We argue that firm heterogeneity is crucial in explaining these different factors. This is especially relevant for the trade-off between choosing a transaction quantity and the deviation of the transfer price from the ALP, which is involved in a firm's profit optimization problem. In this regard, we investigate the effect of firm heterogeneity in terms of productivity, transportation costs, and the difference in corporate income tax rates, on import prices and quantities, utilizing data about the universe of import transactions to a tax haven, Switzerland.

We contribute to previous literature by generalizing the concealment cost function and proposing a theory framework, in which transfer prices deviate from arm's length prices for a tax avoidance motive.[1] Moreover, we introduce firm heterogeneity in terms of productivity in our theoretical model of transfer pricing to examine the determinants of transfer pricing along *extensive* and *intensive* margins of transfer pricing. Next, the model allows for a theory-consistent estimation of concealment cost function parameters, which can be

[1] Concealment costs are the costs that arise due to engaging in transfer pricing, i.e., paying accountants, dealing with tax authorities, etc.

used for counter-factual analysis within our model framework. Analyzing how variations in tax differentials, trade costs, or trade policy instruments affect tax evasion or limit transfer pricing might be an important policy tool. In our empirical analysis, we employ detailed information about the universe of Swiss firm-level manufacturing import transactions between 2006 and 2015. This dataset offers several benefits. Due to low corporate taxes and a favorable economic and institutional environment with generous tax privileges for certain legal forms, Switzerland is a highly attractive location for foreign-owned MNEs. At the same time, Switzerland is a major location for domestically owned MNEs (for an overview see Egger and Koethenbuerger, 2016). Second, we use name matching to separate intra-firm transactions from ALP transactions within firms. Thus, in contrast to the recent literature on transfer pricing focussing on variation across countries, we are able to exploit the variation within a firm to clearly identify mispricing. Third, we are able to exploit both the variation in corporate income tax rates across local jurisdictions in Switzerland and across exporter countries to determine how tax differentials affect profit shifting. The focus on a single importer country allows us to hold pricing-to-market and other general importer conditions constant. We also include transportation costs to explain the behavior of MNEs with regard to transfer pricing. Lastly, quantifying the extent of transfer mispricing helps to understand the direction and pattern of global trade flows, and to which extent variations in corporate taxes drive trade those patterns.

Similar to Davies et al. (2018) we develop a theoretical model in which transfer pricing is explicitly governed by concealment costs, assuming that firms misprice internal transactions because they aim at avoiding taxes. Firms produce differentiated goods and have monopoly power for their variety. They either produce the inputs needed for the differentiated good themselves in the domestic country, or they import them from a foreign affiliate. The latter makes profit shifting possible. They thereby face three restrictions to engage in transfer pricing. First, concealment costs have a fixed component and are increasing in the deviation from the competitive price as well as the traded quantity of goods. Second, as inputs have to be used in the production process, there is an implicit maximum of inputs that can be used without cannibalizing the domestic monopoly profits. Third, high trade costs might render transfer pricing infeasible, especially for small tax differentials. We find that there exists a certain productivity threshold for firms to engage in profit shifting, which is due to the fixed cost component in the concealment cost function and varies between countries due to country-pair-specific iceberg trade costs. While higher fixed costs are intuitively associated with decreased entry, the role of trade costs is more subtle. Firms will only shift profits if

the gains from profit shifting (in terms of net profits) are sufficiently large to compensate for transportation costs. On the one hand, the net gains from transfer pricing depend on the tax differential – the larger it is, the greater are the potential gains. On the other hand, the costs of transfer pricing increase with transportation costs as goods need to be physically traded to be able to exploit the tax differential. This implies that a trade cost-adjusted tax differential matter. Very high levels of trade cost can even become prohibitive for transfer pricing. Thus, the global decline of transportation costs might play a significant role in the rise of transfer pricing. To this extent, the extensive margin in our theoretical model does not solely depend on the uniform fixed cost component of the concealment cost function, but in a bilateral way on the transportation costs of goods.

The proposed theoretical model is consistent with several empirical facts we observe in the data. In line with previous evidence and conditioning on a host of fixed effects that control for factors such as differences in the price and quantity of imports specific to firms as well as products, that an increase in the tax differential with a higher-tax country of origin by one percentage point leads to about 0.4% lower transfer prices of transactions from that origin, relative to arm's length prices of the same importer firm located in Switzerland. The effect on quantities is positive and amounts to about 2.4%. For imports from countries that have lower tax rates, the effect is reversed. We find that the intra-firm transaction price increases by about 0.8% and the quantity decreases by about 5.5%, compared to transactions at arm's length, of a given firm. We also find that more productive MNEs deviate less from the ALP and ship lower quantities than less productive ones, as the tax differential becomes larger. Given the substantial average tax differentials that importers in Switzerland face, the results are quantitatively important. They are also quantitatively and qualitatively robust to a number of robustness checks.

The paper is structured as follows. The next section describes the related literature. Section 3 provides an overview of corporate taxation in Switzerland. Section 4 outlines the theoretical model. In section 5 we describe the data and the empirical estimation strategy. Finally, section 6 concludes.

2 Related Literature

A large strand of the public finance literature focuses on profit shifting (see Huizinga and Laeven, 2008; Schindler and Schjelderup, 2016; Dharmapala, 2014), as well as channels through which MNEs reduce their corporate tax payments. Torslov et al. (2018) estimate

that close to 40% of all MNEs profits are shifted to tax havens each year. Focussing on firm level data, evidence has shown that MNEs avoid taxation by shifting profits to low tax countries (e.g. Clausing, 2009; Dyreng et al., 2012; Dharmapala and Riedel, 2013; Egger et al., 2014). For instance, the results in Egger et al. (2010) suggest that, among European firms, the absolute tax payments of MNEs are lower than those of comparable firms that only operate domestically. The main instruments used for profit shifting are transfer pricing (see Bernard et al., 2006; Davies et al., 2018), debt shifting (see Egger et al., 2014), and royalty payments (see Karkinsky and Riedel, 2012; Griffith et al., 2014 among others). Bernard et al. (2006) and Davies et al. (2018) have shown for the US and France, respectively, that deviations of the transfer price from ALP are related to differences in taxes. The former find that the price wedge depends on product differentiation, firm size, market power, destination-country tax rates, and import tariffs for MNEs in the US, and the latter show that the biggest French MNEs consistently use price transfers.

These findings imply that transfer pricing is an important instrument for multinational firms' tax avoidance practices, and that intensive local multinational activity creates some leeway for profit shifting. The OECD (2010) states the problem underlying the use of transfer prices as a vehicle for profit shifting: *"When independent enterprises deal with each other, the conditions of their commercial and financial relations (e.g., the price of goods transferred or services provided and the conditions of the transfer or provision) ordinarily are determined by market forces. When associated enterprises deal with each other, their commercial and financial relations may not be directly affected by external market forces in the same way."*

Indeed MNEs may use sophisticated methods, for instance, by engaging in manipulation of the arm's length price to conceal deviations in the transfer price (Cristea and Nguyen, 2016) for tax reasons. Even without mispricing intent, Bauer and Langenmayr (2013) argue that MNEs are more productive and hence the marginal cost for an intra-firm transaction is lower than that of an independently sourced input. Additionally, the latter involves a bargaining mark-up, which can make it relatively more expensive. In Keuschnigg and Devereux (2013), financial frictions distort the transfer price. All these factors render the empirical analysis of transfer pricing practices difficult. The next section discusses the institutional setup in Switzerland that fosters profit shifting, in particular transfer pricing.

3 Corporate Taxes and Profit Shifting in Switzerland

The topic of tax-induced profit shifting of MNEs is a highly debated one, having substantial quantitative implications with respect to the potential losses for national tax authorities. For instance, the Australian Administrative Appeals Tribunal decided in 2008 that Roche Products Pty Limited Australia, a subsidiary of Roche Holdings Ltd of Basel, Switzerland had overpaid 45 million dollar for its ethical pharmaceutical products to Roche Basel (Switzerland) between 1993 and 2003.

Switzerland has a residence-based corporate tax system such that companies are subject to corporate income tax on worldwide income, with the exception of income attributable to foreign permanent establishments or foreign immobile property. Low corporate tax rates in general and privileged taxation of holding companies, administrative companies and mixed companies on a cantonal level (similar to US state level) in Switzerland as well as its reputation as a tax haven, provide an incentive for MNEs to use transfer pricing, among other methods, in order to avoid taxes. Although the privileged taxation is supposed to be abolished through a current corporate tax reform to be put in place by 2018 or 2019, the period under consideration in this paper still falls under current practices.[2]

– Table 1 about here –

Table 1 summarizes the 26 Swiss cantonal corporate profit tax rates for incorporated companies in 2015. These data represent the tax burden in percent calculated by the Swiss Tax Administration for each of the 26 cantonal capitals. It is evident that the variation in tax rates is large, and the lowest rates apply in central Switzerland, with an average corporate tax rate of little more than 12 percent in the canton of *Nidwalden*. The highest tax rates apply in Western Switzerland and Basle, amounting to about 25 percent in Geneva. The average tax rate is 18.4 percent in 2015, and 19.5 percent across all years between 2006 and 2015.

Profits are supposed to be taxed in the location in which the value added was generated. In general market forces present in transactions between two independent firms should enforce the aforementioned principle. Thus, transactions between two independent firms should be priced at arm's length. In contrast, multinational firms may exploit a lack of market forces (within the firm) to shift profits to low tax countries by pricing this internal transactions

[2]Tax privileges for MNEs will be substituted by a decrease in corporate tax rates as well as other benefits more in line with the OECD BEPS package. This includes improved bilateral and multilateral cooperation to diminish profit shifting. In addition, bank secrecy for foreign clients will be abolished at around the same time, hence it is questionable whether Switzerland can be considered a tax haven in the future.

different from the ALP. This involves undercharging for inputs sourced from affiliated firms located in high tax countries, and/or overcharging inputs supplied from affiliated firms in low tax countries. The identification of mispriced transactions is thus empirically challenging. Comparable goods may not be available in the data, especially for specialized industries such as the pharmaceutical industry, which is an important sector for the Swiss economy and a highly regulated market, including the regulation of prices.[3] Accordingly it is not peculiar per se, that transfer prices deviate from arm's length prices. The characteristics of institutions in Switzerland and a unique dataset about Swiss trade transactions and firm information, together with a theoretical model laid out in the next section, which establishes a framework for the empirical estimation in the subsequent chapters, allow us to overcome many of the challenges present in previous work. Still Switzerland is a small open economy with a diverse import and export structure, which allows us to investigate transactions on a highly disaggregated level. Moreover, Switzerland is one of a few tax havens that is actually capable of processing larger amounts of imports, which is very relevant for transfer pricing of physical good transactions. Last, but not least, the Swiss customs data is of high quality.

4 Model

In this section we provide a model of transfer pricing between two affiliated firms located in two generic countries indicated by $i = d, f$, where d indicates the domestic country and f the foreign country. Without loss of generality we write the model from the perspective of a firm in the domestic country that potentially has affiliates in the foreign country. Both countries differ in their corporate tax rates, $\tau_i \in [0, 1]$, which will create incentives to shift profits between the two affiliated entities. We assume that firms are heterogeneous in terms of their productivity, ϕ, which is drawn from a known distribution. Firms in the domestic country produce differentiated goods indexed by ω. For simplicity we assume that the differentiated goods can only be produced and consumed locally in the domestic country, but firms are able to use inputs either produced by domestic affiliate and/or import inputs provided by a foreign affiliate. The foreign affiliate does not produce any differentiated goods and its sole purpose is to decrease the effective tax rate of the domestic firm.[4] We denote inputs provided by the domestic entity by x_d and foreign inputs by x_f. Both inputs are perfect

[3] The Swiss pharmaceutical industry contributes about 6% to the Swiss GDP and accounts for about 30% of exports.

[4] Allowing for foreign production and even trade in differentiated goods, as in Helpman et al. (2004) does not change the analysis, but greatly increases the model complexity as it introduces further entry cutoffs and makes aggregation more difficult.

substitutes in the production process of the differentiated goods. Profit shifting by transfer pricing arises from the fact that an input can be imported from a foreign affiliated firm and its price can be set differently from the market price.

Concealment costs

If a firm decides to source inputs from a foreign affiliated firm, it may engage in transfer mispricing, i.e., shifting profits towards the country with the lower corporate tax rate. As in Davies et al. (2018) firms can shift profits towards low tax countries by exporting (importing) intermediate inputs above (below) the market price. As this kind of mispricing usually violates tax laws, the firm incurs concealment costs. These costs might include paying transfer pricing accountants, preparing documentation, and legal fees.

Depending on the tax differentials firms either want to import goods below or above the arms' length price. Thus, we specify two generalized concealment cost functions: (i) one in which the foreign firm undercharges their inputs, i.e., the corporate tax rate in the domestic country is lower than in the foreign country, and (ii) one for a foreign firms that overcharges, i.e., the foreign corporate tax rate is lower than the domestic one.

In our model, concealment costs consist of two parts: a variable part and a fixed part. The firm has to pay some fixed cost, F_U and F_O, if the firm uses undercharging or overcharging, respectively. The variable part depends positively on the difference between the transfer price and the competitive price, and the total amount of goods shipped. If the transfer price and the competitive price differ significantly, the firm may be audited by the tax authority and needs to justify this difference, which increases costs. If a firm exports large quantities, this can imply greater profit shifting and tax authorities might demand more documentation, i.e., greater exports/imports might raise the suspicion of tax authorities.[5] The concealment costs for undercharging transfer pricing are

$$(p_A - p_U)^{\alpha} x_f^{\beta} + F_U, \tag{1}$$

where p_U is the undercharging price of an internal transaction, which is lower than the competitive price p_A, and $\beta, \alpha \geq 0$. If the firm overcharges, i.e., the transfer price of input

[5] In theory greater quantities might lower the concealment costs due to economies of scale, i.e., in out model the β parameter could be negative. We neglect this possibility as the empirical evidence later show that $\beta > 0$.

p_O is higher than the competitive market price, the concealment costs are

$$(p_O - p_A)^\alpha x_f^\beta + F_O. \tag{2}$$

In contrast to Cristea and Nguyen (2016) firms take the competitive arms' length price as given. Thus, instead of strategically setting p_A and p_U (p_O) to reduce concealment costs, the firm only sets the price wedge between the arms' length price and the internal transfer price p_U or p_O by choosing the optimal p_U and p_O, respectively.[6]

Utility

Due to our model setup it is sufficient to consider only the demand for the differentiated good in the domestic country. We assume that individuals (in the domestic country) have linear-quadratic preferences as in Melitz and Ottaviano (2008):

$$U = q_0 + e\int_{\omega\in\Omega} q(\omega)d\omega - \frac{1}{2}b\left(\int_{\omega\in\Omega} q(\omega)d\omega\right)^2 - \frac{1}{2}c\int_{\omega\in\Omega} q(\omega)^2 d\omega, \tag{3}$$

where q_0 is a numeraire, $q(\omega)$ is the quantity of a differentiated good ω from a set Ω of available goods in the domestic country, and e, b, and c are positive constants. The aggregate market demand is

$$p(\omega) = e - b\int_{\omega\in\Omega} q(\omega)d\omega - cq(\omega), \tag{4}$$

where we assume that the mass of individuals in the economy is normalized to one. We define the aggregate price of all products as

$$P \equiv \int_{\omega\in\Omega} p(\omega)d\omega, \tag{5}$$

and hence

$$P \equiv \left(e - b\int_{\omega\in\Omega} q(\omega)d\omega\right)M - c\int_{\omega\in\Omega} q(\omega)d\omega, \tag{6}$$

where $M \equiv \int_{\omega\in\Omega} d\omega$ is equal to the total number of products in the economy. Then we can write $\int_{\omega\in\Omega} q(\omega)d\omega = \frac{eM-P}{c+bM}$, and the (aggregate) demand for a product ω is linear and given by

$$p(\omega) = E - cq(\omega), \tag{7}$$

[6] In the Appendix A we show that very similar concealment cost functions can be derived using an optimal acting tax authority that maximizes expected fines.

where $E \equiv \frac{ec+bP}{c+bM}$. Without loss of generality we normalize c to one.

Firms

Firms face monopolistic competition in their goods market as in standard Krugman (1980) and Melitz (2003) models. Each differentiated good ω is produced using domestic inputs x_d and/or foreign inputs x_f, which are perfect substitutes in the production process. All domestic parent firms have the same technology to produce the differentiated goods using the two inputs:

$$y(\omega) = x_f + x_d. \tag{8}$$

Firms are heterogeneous in terms of the productivity to produce inputs, x_d and x_f. Their input productivity, ϕ, is drawn from a known distribution with cumulative density function $G(\phi)$. We assume that firms can produce one unit of input x_d and x_f, respectively at costs $1/\phi$ in both countries. As the productivity of foreign affiliates and the domestic parent firm is the same, there is no reason for firms to source any input from abroad. Thus, in this stylized model input sourcing from abroad will be exclusively driven by tax considerations.[7]

In our basic model two types of firms exists. First, a firm that only sources domestically. Second, a firm that sources from a foreign affiliate, but at a price below (above) the arm's length price, i.e., it uses transfer mispricing to shift profits to locations with lower tax rates.

Domestic sourcing firm

Net profits of a firm that only operates domestically, sources its inputs domestically, and sells its differentiated good in the domestic market, are:

$$\pi^d(\omega) = (1 - \tau_d)\left(p(\omega)y(\omega) - \frac{x_d}{\phi}\right) - F, \tag{9}$$

where we use that for an exclusively domestically sourcing firm, $y(\omega) = x_d$. F denotes fixed

[7] We could allow firms to have two independent draws from the $G(\phi)$ distribution, the first one for the domestic productivity, ϕ_d, the second one for the foreign productivity ϕ_f. Thus, $G(\phi_d > \frac{\phi_f}{\epsilon})$ gives the probability of domestic sourcing without the possibility of transfer pricing, where $\epsilon \geq 1$ are iceberg trade costs. Introducing transfer pricing will alter this relationship. The tax differential between the domestic and foreign country might render imported inputs more profitable. Thus, the foreign productivity draw necessary for importing might be lower, and everything else equal more firms will import from foreign affiliates. Or the fixed costs component of the concealment cost function are so high that $\frac{\phi}{\epsilon} < \hat{\phi}$, where $\hat{\phi}$ is necessary productivity of a firm to use transfer pricing. In this case we have three types of firm: (i) firms that only source domestically, with the lowest productivity, (ii) firms that source from abroad but at ALP, with intermediate productivity, and (iii) firms that import from the foreign country at a price lower (higher) than the arm's length price. This would introduce more cutoffs for firms without adding much more insights.

setup costs. The (internal) costs of the firm for its inputs is $1/\phi$. Because each firm has a monopoly in its market, the demand for inputs x_d is given by

$$x_d = \frac{E\phi - 1}{2\phi}, \tag{10}$$

which is increasing in ϕ. Then the necessary productivity to cover the fixed setup cost F and to enter the domestic market is given by

$$\overline{\phi} \geq \left(E - 2\sqrt{\frac{F}{1-\tau_d}} \right)^{-1}. \tag{11}$$

Undercharging MNE

Assuming that the domestic tax rate is lower than the foreign tax rate, $\tau_d < \tau_f$. In this case firms want to shift profits from the foreign high-tax country towards the domestic low-tax country. Thus, the firm will import intermediate inputs below the arms' length price, i.e., undercharge. Total consolidated (foreign and domestic) net profits of the firm are:

$$\begin{aligned} \pi^U(\omega) =& (1-\tau_d)\left(p(\omega)y(\omega) - \frac{x_d}{\phi} - p_U \epsilon x_f \right) \\ &+ (1-\tau_f)\left(p_U - \frac{1}{\phi} \right) x_f - (1-\tau_f)(p_A - p_U)^\alpha x_f^\beta - F_U - F, \end{aligned} \tag{12}$$

where gross profits generated by the domestic entity are revenues less domestically produced inputs and *undercharged* foreign input costs. The gross profits (losses) of the foreign entity are the revenues from the undercharged exports less the production costs. The term $(p_A - p_U)^\alpha x_f^\beta - F_U$ reflects the concealment costs, which further decrease consolidated profits. Note that the fixed setup costs, F, are not deductible. We assume that the concealment costs are paid by the entity with the higher tax rate to even further decrease profits in the high-tax location.[8] If a firm wants to use foreign imports (to make use of transfer pricing), they face iceberg trade costs $\epsilon \geq 1$. These trade costs imply a constraint to shift profits, i.e., the net profit gains of transfer pricing must be sufficiently high to pay for the fixed costs, F_U, and the inefficiency created by the presence of iceberg trade costs.

We use the fact that firms have monopoly power and that $y(\omega) = x_d + x_f$, to derive the first

[8] Tax authorities in high-tax countries are more likely to investigate shifted profits, while low-tax countries actually gain tax revenues from the profit shifting. Thus, it makes sense to assume that the documentation and legal fees occur in the high-tax country.

order conditions of the firm's maximization problem:

$$\begin{aligned}
\frac{\partial \pi(\omega)^U}{\partial x_d} &= 0 \quad \rightarrow E - 2(x_d + x_f) - \frac{1}{\phi} = 0, \\
\frac{\partial \pi(\omega)^U}{\partial x_f} &= 0 \quad \rightarrow (1-\tau_d)\left(E - 2(x_d + x_f) - p_U \epsilon\right) + (1-\tau_f)\left(p_U - \frac{1}{\phi}\right) - (1-\tau_f)(p_A - p_U)^{\alpha} \beta x_f^{\beta-1} = 0, \\
\frac{\partial \pi(\omega)^U}{\partial p_U} &= 0 \quad \rightarrow -(1-\tau_d)\epsilon x_f + (1-\tau_f) x_f + (1-\tau_f)\alpha (p_A - p_U)^{\alpha-1} x_f^{\beta} = 0.
\end{aligned}$$

The optimal undercharging price, p_U, is given by

$$p_U = \frac{\alpha}{\beta - \alpha}\left(\frac{\beta}{\alpha} p_A - \frac{\tau_f - \tau_d}{\phi \zeta}\right), \tag{13}$$

where $\zeta = (1-\tau_d)\epsilon - (1-\tau_f)$ is the trade costs adjusted tax differential. Note that if $\tau_d < \tau_f$ and $\epsilon \geq 1$, then $\zeta > 0$.

In order to shift profits, the undercharging price needs to satisfy $0 \leq p_U \leq \frac{1}{\phi}$. The first inequality rules out negative prices, the second inequality states that the undercharging price must be smaller than the cost of the affiliate, $\frac{1}{\phi}$, otherwise no profits would been shifted. While the second inequality always holds, given that $p_A = \frac{1}{\phi}$, $\epsilon \geq 1$ and $\alpha > \beta$, the first inequality constrains firms' pricing behavior. If $p_A = \frac{1}{\phi}$ and $\alpha > \beta$, we need $\zeta < \frac{\alpha}{\beta}(\tau_f - \tau_d)$ to ensure that $0 \leq p_U$. Solving the inequality for ϵ yields:

$$\epsilon \leq \frac{\alpha}{\beta} \frac{\tau_f - \tau_d}{1 - \tau_d} + \frac{1 - \tau_f}{1 - \tau_d}, \tag{14}$$

where the right hand side (RHS) is always greater than 1. Notice that the RHS is increasing in τ_d and decreasing in τ_f, i.e., increasing the tax differential between the two countries makes undercharging more feasible. Intuitively, if transportation costs are too high relative to the tax differential, the firm wants to set a negative price to be able to shift profits, i.e., the exporter is actually paying the importer for the intermediate goods. We assume that this extreme case of transfer pricing is not possible. Figure 1 shows the graphical relationship described in equation (14). For too high ϵ no transfer pricing is feasible, and the threshold increases with the tax differential, $\delta\tau$.

– Figure 1 about here –

We assume for the moment that E is sufficiently large to ensure that x_d and x_f are non-negative in equilibrium, i.e., we are in an interior solution in which firms use both foreign and domestic inputs. Additionally, assuming that ϵ is sufficiently small, such that for a firm

with given productivity, ϕ, the undercharging price is positive, and $\alpha > \beta$, we can show that

$$\partial(p_A - p_U)/\partial\phi < 0, \qquad \partial(p_A - p_U)/\partial\epsilon > 0,$$

$$\partial(p_A - p_U)/\partial\tau_f > 0, \qquad \partial(p_A - p_U)/\partial\tau_d < 0.$$

Thus, more productive firms charge lower prices, but deviate less from the competitive price. On the other hand firm deviate more if the tax differential is greater, holding productivity constant. Higher trade costs increase the price wedge. Importantly, the undercharging price actually declines with productivity, i.e.,

$$\partial p_U/\partial\phi < 0.$$

This stands in sharp contrast to the finding that the price wedge declines with productivity. As firms become more productive, their arm's length price declines, given that $p_A = 1/\phi$. Still the undercharging price declines less relative to p_A.

Similarly, we derive the optimal traded quantity as

$$x_f^U = \left(\frac{(1-\tau_d)\epsilon - (1-\tau_f)}{(1-\tau_f)\alpha(p_A - p_U)^{\alpha-1}}\right)^{\frac{1}{\beta-1}}. \tag{15}$$

Substituting the optimal undercharging price, p_U, from equation (13) yields

$$x_f^U = \left(\frac{(1-\tau_d)\epsilon - (1-\tau_f)}{(1-\tau_f)\alpha\left(\frac{\alpha}{\alpha-\beta}\left(p_A - \frac{\tau_f-\tau_d}{\phi\zeta}\right)\right)^{\alpha-1}}\right)^{\frac{1}{\beta-1}}. \tag{16}$$

Assuming that $p_A = \frac{1}{\phi}$, an interior solution, $(x_f, x_d > 0)$, and $\alpha > \beta > 1$, it can be shown that $\partial x_f^U/\partial\phi > 0$. Thus, bigger (more productive) firms trade more goods. Moreover, we find that

$$\partial x_f^U/\partial\tau_f > 0, \qquad \partial x_f^U/\partial\tau_d < 0, \qquad \partial x_f^U/\partial\epsilon < 0.$$

Every thing else equal, a higher foreign tax rate (higher tax differential) incentivizes a firm trade more goods to shift profits. Moreover firm deviate more from the competitive price, thus total profits shifted increase with the tax differential. Given that $\alpha > \beta$ and an interior solution, more productive firms will deviate less from the arms' length price and instead use more quantity. With increasing productivity firms substitute price deviation for quantity.

Overcharging MNE

In the Appendix B we explicitly derive the analogous expression for an overcharging MNE. The overcharging price is given by

$$p_O = \frac{\beta}{\beta - \alpha}\left(p_A + \frac{\alpha}{\beta}\frac{\tau_d - \tau_f}{\phi\zeta}\right). \tag{17}$$

Note that the overcharging price has to be higher than the arms' length price, which is the case if $\zeta \leq 0$. Solving for ϵ we get an analogous condition to equation (14) for the feasibility of profit shifting in the presence of iceberg trade costs:

$$\epsilon < \frac{1 - \tau_f}{1 - \tau_d}, \tag{18}$$

where the RHS is greater than one as $\tau_d > \tau_f$. In Figure 1 the left side shows the relationship given by equation (18). Similarly, ϵ must be sufficiently small to make overcharging feasible and this threshold increases with the absolute tax differential.

The optimal quantity in the overcharging case is

$$x_f^O = \left(\frac{(1 - \tau_f) - \epsilon(1 - \tau_d)}{(1 - \tau_d)\alpha(p_O - p_A)^{\alpha - 1}}\right)^{\frac{1}{\beta - 1}}. \tag{19}$$

Binding non-negativity constraint

Also the assumption that E is sufficiently large to ensure that $x_d, x_f > 0$ might seem trivial, yet it has important implications for the model. If the domestic demand for firm's differentiated good is small, only limited amounts of inputs can be used in the production. This limits the profit shifting capabilities of a firm. If profit shifting is optimal and E is not sufficiently large, we are in a corner solution with $x_d = 0$. In this case the firm will source all its inputs from abroad, still maximizing local profits, i.e., the firm behaves like a monopolist in the domestic market and does not cannibalize its monopoly profits. Thus, the optimal inputs with a binding non-negativity constraint on x_d are given by

$$\tilde{x}_f = \left(\frac{\phi E - \epsilon}{2\phi}\right). \tag{20}$$

The first order condition becomes

$$\frac{\partial \pi(\omega)^U}{\partial p_U} = 0 \quad \rightarrow -(1 - \tau_d)\epsilon + (1 - \tau_f) + (1 - \tau_f)\alpha(p_A - p_U)^{\alpha - 1}\hat{x}_f^{\beta - 1} = 0,$$

which implies that $\frac{\partial(p_A - p_U)}{\partial \phi} > 0$. More productive firms deviate more from the arms' length price. The only possibility to shift more profits is to change the price wedge, e.g., lower p_U. Moreover, for constraint firms we will not observe any quantity adjustments, as they always use the maximum amounts of inputs. Thus, firm's with a binding non-negativity constraint will introduce a downward bias in the price difference estimation and in the quantity estimation. To this extend our estimates will be lower bounds.

Extensive margin

Firms face two constraints when using transfer pricing for reasons of tax avoidance. First, the transfer price has to satisfy the two conditions given by equations (14) and (18), depicted in Figure 1. Only if the iceberg trade costs are below the line, transfer pricing (over- or undercharging) will be feasible. The part to the left of the kink corresponds to the overcharging case, i.e., $\tau_f \leq \tau_d$, while the right-hand side gives the undercharging condition. The depicted relationship is independent of ϕ, as we assume the arm's length price to be equal to $\frac{1}{\phi}$. Clearly, transfer mispricing is more feasible the greater tax differential, as the upper limit of trade cost increases in the right and left tail.

Second, firms have to recover the fixed costs. Specifically, in the case of undercharging the additional profits of transfer pricing less the concealment costs and total trade costs must exceed the fixed costs, F_U. Equation (21) states this relationship:

$$(\tau_f - \tau_d)x_f^U(p_A - P_U) - (1 - \tau_f)(p_A - p_U)^\alpha (x_U^U)^\beta - (1 - \tau_f)(\epsilon - 1)x_f^U p_U \geq F_U, \qquad (21)$$

where the first part of the left-hand side (LHS) gives the additional (net) profits due to profit shifting, the middle part represents the reduction of (net) profits due to the variable costs of undercharging, and the last part are the (net) additional iceberg transportation costs a firm faces. Moreover, the LHS increases with ϕ, thus there is a cutoff value $\hat{\phi}$ for which firms will start to use transfer pricing. i.e., the above inequality holds with equality. We substitute the optimal undercharging price and import quantities from equation (13) and (16), respectively. Equation (21) cannot be solved explicitly for $\hat{\phi}$. Using the implicit function theorem, we find that $\hat{\phi}$, decreases with the tax differential and increases with transportation costs, as well as fixed costs of transfer pricing.

Corollary

The specific concealment cost function in the model yields some interesting and testable

predictions. Given an interior solution, $\alpha > \beta > 1$, and trade costs sufficiently low the following model predictions arise:

1. The necessary productivity of firms to engage in transfer mispricing increases with the fixed costs of transfer mispricing and per piece trade costs, and decreases with the tax differential.

2. The optimal quantity shipped of transfer mispriced goods depends negatively on the iceberg trade costs, and positively on the tax differential.

3. The optimal price wedge is inversely related to the quantity traded, thus more goods are traded at prices more similar to the arm's length price.

5 Empirical analysis

5.1 Data

We combine several micro-level and aggregate data for Switzerland. First, we use the universe of Swiss import transactions in the manufacturing sector between 2006 and 2015, provided by the *Swiss Federal Customs Administration EZV (Oberzolldirektion)*.[9] These contain information about the CIF (including cost, insurance and freight) transaction volumes and quantities, the 8-digit product category (*Tarifnummer*) based on the HS classification, the country of origin, and the name of both the importer and the exporter, as well as the location at the level of zipcodes of the importer. The availability of volumes and quantities allows us to construct CIF unit values, which serve as the proxy for the transaction-specific mark-up. Note that firms are not identified by a unique identification number but only based on name and location characteristics, hence a string search algorithm was applied in order to identify firms.[10] Next, Stata's *reclink2* command was used to identify pair-relationships, i.e., affiliates or parent companies based on the name and country of origin of the importer or exporter, respectively. As not all affiliated or holding companies carry the same name as the Swiss firm, this results in conservative estimates throughout the next section. Second, we have matched these data to information about firm-level characteristics such as

[9] We abstain from using export data, because the data about export transactions represent only a fraction of export sales. Furthermore, they would not allow us to hold characteristics related to Switzerland constant.

[10] We thereby treated firms with several locations in Switzerland as separate entities. The search process was complicated by an undefined number of firm and location spellings, and it was computationally expensive because of the large sample size. We have reduced the burden by conditioning on the canton for Swiss firms and on the country of origin for foreign firms, and by matching strings that shared the same first character. The search was implemented by using the Stata plugin *strgroup*.

the number of employees, the NACE affiliation, operating revenue and capital stock from Bureau van Dijk's *Amadeus* database.[11] These data are used to calculate log productivity based on Petrin et al. (2004). The free variable is the log number of employees, the proxy we use is the firm's total import volume, capital is taken from the balance sheet, and we use gross operating turnover as the dependent variable. We also deflate the variables using the producer price indices from the production accounts (national accounts) of the Swiss statistical office (BFS), of the respective industry the firm is affiliated with.[12] Note that this substantially decreases the number of available observations, for which firm-level information is available. Third, we use cantonal corporate profit tax rates, which are matched to the canton the importer is located in. These data are summarized in Table 1 and described in Section 3.[13] These may represent upper bounds, as the tax rates for holdings summarized in Appendix Table 6 show. We will account for this in Section 5.4 accordingly. Fourth, we use comprehensive data for corporate profit tax rates in 79 countries between 1996 and 2013 as described in Egger et al. (2015). For the corporate profit tax rate we use the maximum corporate profit tax rate in a country and year.

– Figure 2 about here –

Figure 2 shows the distribution of the aforementioned tax rate for each year in the data, using whisker-plots. The area around the median (a horizontal bar) indicated by a box refers to the interquartile range (IQR), whereas the extended lines, the *whiskers*, indicate values within a maximum of 1.5 times the IQR. The corporate profit tax rates in Figure 2 show a relatively high degree of variability over time, even at the median.

– Table 2 about here –

Next, we provide summary statistics about the data used for empirical analysis in Table 2.[14] Panel A of the table provides statistics about tax differentials. We observe nearly 22 thousand canton-origin-year pairs with positive tax differentials. This represents the case where the foreign tax exceeds the tax of a Swiss canton. The reverse case covers roughly 8.7 canton-origin-year pairs. The average overall tax differential amounts to 4 percentage points,

[11] Alternatively, we have tried to match this sample with information about firm ownership from Bureau van Dijk's *Orbis* database, however this resulted in a very small number of matches (around 50).

[12] We have calculated productivities differently (according to the Ackerberg et al. (2015) and Gandhi et al. (2016) methods, as well as turnover over the number of employees) to check the sensitivity of the results. The corresponding results were unchanged.

[13] We use corporate profit tax rates for incorporated companies with capital and reserves of 2 million Swiss Francs. The tax rates for ones with capital and reserves of 100 thousand Swiss Francs are basically unchanged.

[14] Note that the high sensitivity of the data do not allow us to illustrate summary statistics at more disaggregated levels.

whereas it is about 9 percentage points for pairs with $I_{\Delta\tau_{fct}>0}$ and about 8 percentage points (in absolute terms) for pairs with $I_{\Delta\tau_{fct}<0}$. The corresponding maximum tax differentials amount to 40 and 27 percentage points, respectively

Summary statistics about import transactions are reported in Panel B. We have collapsed the data at the level of the firm, the exporter firm, the country of origin, product code (by HS 8-digit industry), as well as year.[15] This reduces computational requirements as well as the noise inherent in an analysis based on the use of every single import transaction as recorded by the customs office. The resulting number of observations amounts to nearly 21 million. Within these, we observe close to 400 thousand firms, importing 8.4 thousand products from 1.7 million exporter firms located in 189 countries, in 10 years. A small proportion of observations refers to intra-firm transactions (about 135 thousand). This is well in line with patterns for other countries, but might also reflect in part the fact that we are only able to determine firm relationships based on exporter and importer names, hence if an affiliate carries a different firm name, this will be classified as an independent transaction in our data. The log price and log quantity are similar, relatively widely dispersed, and amount to 3.5 on average. A further decomposition shows that intra-firm transactions have a higher price unconditionally, and the quantities are also larger, irrespective of whether they stem from a country that offers higher or lower taxes than a given Swiss canton.

Panel C illustrates statistics about log productivity as well as the number of observations in the sample for which we can calculate this measure. We are able to do so for a little more than 10% of the previous data.

5.2 Reduced form estimation

The first part of the empirical strategy is related to Davies et al. (2018) and Liu et al. (2017). Precisely, we are interested in testing the predictions obtained in section 4 by way of reduced form regressions. As described in the previous section, the unit of observation is an import transaction, specific to the import of a good in product category k by importer firm i (located in canton c) from exporting firm j (located in origin country f) in year t.[16] The notation illustrates that the data used is novel in its degree of detail, as we can observe import pair-relationships at the product level, and exploit variation in regional taxes in addition to the one in countries of origin. The price baseline regression is estimated as follows:

[15] We have tested the statistical patterns and the robustness of our empirical results to using monthly data instead. This did not change the results in substantial ways. The results are reported in Appendix D.

[16] Because i is specific to c, and j to f, we simplify the notation accordingly.

$$
\begin{aligned}
p_{\text{ijkt}} = {} & \gamma_1(I_{\Delta\tau_{\text{fct}}>0} \times \Delta\tau_{\text{fct}}) + \gamma_2(I_{\Delta\tau_{\text{fct}}>0} \times \text{MNE}_{\text{ijt}} \times \Delta\tau_{\text{fct}}) + \gamma_3(I_{\Delta\tau_{\text{fct}}<0} \times \Delta\tau_{\text{fct}}) \\
& + \gamma_4(I_{\Delta\tau_{\text{fct}}<0} \times \text{MNE}_{\text{ijt}} \times \Delta\tau_{\text{fct}}) + \kappa_{\text{i(c)kf}} + \lambda_{\text{ikt}} + \mu_{\text{fkt}} + \varepsilon_{\text{ijkt}},
\end{aligned} \tag{22}
$$

where p_{ijkt} is the log import price for each transaction of good k between firm i located in a Swiss canton c, and a foreign firm j located in a country f, $I_{\Delta\tau_{\text{fct}}>0}$ equals one if f's tax rate is higher than c's in t, $I_{\Delta\tau_{\text{fct}}<0}$ equals one if f's tax rate is lower than c's in t, MNE_{ijt} is a binary variable which equals one if the transaction is intra-firm and zero if it is an arm's length transaction, $\Delta\tau_{\text{fct}} \equiv \tau_{\text{ft}} - \tau_{\text{ct}}$ measures the corporate tax rate differential between country of origin f and canton c in t, and $\varepsilon_{\text{ijkt}}$ is the disturbance term. The fixed effects $\kappa_{\text{i(c)kf}}$, λ_{ikt}, and μ_{fkt} remove the bias from unobserved factors that affect prices (e.g., time-invariant firm-product-origin factors accounting for trade costs such as distance, common language; time-varying factors specific to the firm-product tuple such as productivity, quality and size; and time-varying supply and demand shifters common to firms and cantons, such as tariff and nontariff barriers, and most importantly, average product prices which we assume are capturing appropriate ALP's within sufficiently detailed product categories). The fixed effects also account for the fact that multinational companies' prices may have generally lower levels (e.g., Bauer and Langenmayr, 2013). Essentially we are using the variation of different transactions within a firm for a given product, origin and year to identify the effect of differences in taxes on transactions with affiliates versus transactions at arm's length, on the price wedge. We cluster the standard errors by canton-origin-time to account for serial correlation and correlation within canton-origin tuples.
The equation for log import quantities and volumes can be written analogously to (22).

Next we can adapt estimation of equation (22) to be based on the data containing firm-level information from Amadeus. This leads to an equation which differs from (22) in the inclusion of log productivity and takes the following form:

$$
\begin{aligned}
p_{\text{ijkt}} = {} & \delta_1(\text{MNE}_{\text{ijt}} \times \phi_{\text{it}}) + \delta_2(I_{\Delta\tau_{\text{fct}}>0} \times \text{MNE}_{\text{ijt}} \times \Delta\tau_{\text{fct}}) \\
& + \delta_3(I_{\Delta\tau_{\text{fct}}>0} \times \text{MNE}_{\text{ijt}} \times \Delta\tau_{\text{fct}} \times \phi_{\text{it}}) + \delta_4(I_{\Delta\tau_{\text{fct}}<0} \times \text{MNE}_{\text{ijt}} \times \Delta\tau_{\text{fct}}) \\
& + \delta_5(I_{\Delta\tau_{\text{fct}}<0} \times \text{MNE}_{\text{ijt}} \times \Delta\tau_{\text{fct}} \times \phi_{\text{it}}) + \eta_{\text{i(c)kf}} + \xi_{\text{fkt}} + \nu_{\text{ijkt}},
\end{aligned} \tag{23}
$$

where ϕ_{it} is log productivity, and ν_{ijkt} is the disturbance term. This equation includes the interaction terms as predicted by the model in Section 4. These include the effect of produc-

tivity for MNE transactions, the effect of the tax difference for MNE transactions, and the triple interaction term of MNE_{ijt}, the tax wedge, and productivity. In contrast to eq. (22), eq. (23) has to drop the fixed effects in the *ikt* dimension, as these are collinear with ϕ_{it}. The equation for log quantity can be stated analogously.

5.3 Results

This section is structured as follows. We first report results from the estimation of eq. (22) in Table 3. Next, we summarize results from estimating eq. (23) in Table 4. The results from Table 3 allow us to proceed with an estimation of the structural parameters of eq. (24) in the next section. With the results from Table 4 at hand we can verify the predictions derived in section 4, about firm-level factors that determine intensive and extensive margin transfer-pricing patterns.

– Table 3 here –

Panel A of Table 3 reports estimates for log prices, and panel B for log quantities. Columns 1–3 and 5 condition on fixed effects in the *ifk*, *fkt* and *ikt* dimensions, whereas column 4 replaces *ifk* by *icfk*. In column 1 we estimate the effect $\Delta\tau_{\text{fct}}$ for intra-firm transactions only, while we also estimate the effects for transactions at arm's length in the remaining columns. Column 3 reports results for firms that have a least one intra-firm import transaction. Column 5 reports results excluding imports of a number of products, including pharmaceutical products – for instance, pricing of pharmaceutical products is heavily regulated, which might distort our findings – and weapons.[17] The corresponding results imply that for $\tau_{\text{f}} > \tau_{\text{c}}$ and hence $\Delta\tau_{\text{fct}} > 0$, a one percentage point increase in the tax wedge decreases the log price by 0.4%–0.7%, while it increases the log price by 0.8%–0.9% for $\Delta\tau_{\text{fct}} < 0$. The effects are mostly insignificant at conventional levels for imports priced at arm's length. This is expected as the price with an independent party should not be affected by taxes. The analog effects for log quantity as the dependent variable are equally significant statistically, and quantitatively bigger, amounting to 2.4% for $\Delta\tau_{\text{fct}} > 0$ and -5.3% to -5.6% for $\Delta\tau_{\text{fct}} < 0$.

It should be noted that the estimates represent conservative estimates, because we are only able to identify within-firm transactions based on importer and exporter names. If these names differ, the transaction will be classified as an independent one. Yet still, the results

[17] We include only products falling under tariff codes (*Zolltarife* 39-92, 94, 95 and 96. Swiss tariff codes are based on the HS classification.

overall suggest that firms misprice their intra-firm transfers, and they do so more intensively as the tax differential becomes larger. They also ship larger quantities from high-tax countries and lower quantities from low-tax ones as the tax wedge increases. Given an average tax differential of 9 percentage points with higher-tax countries, the price should thus decrease by 3.6% and the quantity should increase by 21.6%. With an average price of roughly 33 CHF and an average quantity of about 35 (in kilograms), this implies that the average intra-firm transaction price is lower by about 1.2 CHF, and the average quantity is higher by around 7.6 kilo. For low-tax countries (with an absolute average tax differential of 8 percentage points), the results suggest that the transaction price is higher by about 2.1 CHF and the quantity decreases by 15.4 kilos.

– Table 4 –

Table 4 reports results corresponding to the estimation of eq. (23), on a subsample of firms for which we observe information necessary to calculate productivity. Again, panel A illustrates results for log price as the dependent variable, and panel B shows those for log quantity. All columns include coefficients for both the first- and second-order effects for MNE transactions as derived in Section 4, by including interaction effects between the MNE dummy, the tax differential, and log productivity. Thereby, the coefficients represent the simple intra-firm slopes. This is in line with theory, which assumes that due to trade costs, the only purpose of shipping inputs is tax avoidance. Column 1 estimates the baseline regression, whereas column 2 focuses on firms with at least on intra-firm transaction, and column 3 replaces the *ifk* with *icfk* fixed effects. Columns 4 and 5 report results based on separate (i.e., non-pooled) regressions for the high- and low-tax cases, respectively, and column 6 excludes products equivalent to column 5 of Table 3. Note that due to the inclusion of productivity, the fixed effects varying in the *ikt* dimension cannot be included.

The results fully square with the theory predictions. We find a negative general productivity effect for intra-firm transaction prices, yet this effect is insignificant except when focusing solely on imports from high-tax countries. The sign of the coefficients is positive regarding log quantities, where also the coefficients are statistically significant and amount to 0.5%–6.7%. Again, the effect of the tax differential in panel A carries a negative sign, implying that a one percentage point increase in the tax wedge decreases the log price by 2.1%–2.3% for $\Delta\tau_{\text{fct}} > 0$, while it increases the log price for $\Delta\tau_{\text{fct}} < 0$, however the latter effects are insignificant. The reverse holds for log quantities as illustrated in panel B, implying that a one percentage point increase in the tax wedge augments the log quantity by 13.2%–13.5%

for $\Delta\tau_{\text{fct}} > 0$. Note that these findings are quantitatively larger than the ones shown in Table 3, which might be due to both sample selection and the exclusion of the *ikt*-specific fixed effects. Finally, we find that the coefficients on the triple interaction term have a positive sign regarding log prices, and a negative sign regarding log quantities for $\Delta\tau_{\text{fct}} > 0$. Thus, more productive firms deviate less from the arms' length prices and ship smaller quantities. The coefficients amount to 0.002 to 0.003 and -0.022 to -0.023, respectively, and are significant statistically. This is particularly striking as the identification is complicated by the limited variation of the variable ϕ_{it} over time. In contrast, this is not the case for $\Delta\tau_{\text{fct}} < 0$. One reason for this might be a greater focus of tax authorities on bigger (more productive) MNEs trading with affiliates located in countries with big tax differentials. This would raises their concealment costs relative to smaller firms and prevent big MNEs from using excessive transfer pricing.

Overall our results imply that more productive firms' intra-firm transaction prices deviate less from their independent transactions, and they also ship lower quantities from affiliates, which is in line with concealment costs (and also trade costs), making profit shifting costly.

– Figures 3 and 4 about here –

The interaction effects are best shown by way of a graphical illustration. We plot the tax wedge against the linear prediction for log prices (LHS) and log quantity (RHS) in Figures 3 and 4.

Figure 3 confirms that the price of import transactions decreases with the tax differential, but only significantly so for intra-firm transactions. Furthermore, the confidence intervals of intra-firm versus arm's length transactions do not overlap. The corresponding pattern is similar yet with a positive sign regarding quantities, where the upward slope for intra-firm transactions is even steeper than for log prices.

Figure 4 illustrates the effect of taxes for intra-firm transactions, evaluated at high versus low productivity, which is measured at the mean +/- one standard deviation. Note that we focus only on the case where the foreign tax is higher ($\Delta\tau_{\text{fct}} > 0$), as the results for proved insignificant in Table 4. The left-hand side of the figure (log prices) illustrate that the confidence intervals around the respective slopes are significant over the entire range of the distribution of the tax differential only for low-productivity firms. For those, the slope is also negative as opposed to the one of high-productivity firms, and towards the middle of the distribution of tax differentials, the confidence intervals do not overlap. The right-hand side of the figure (log quantity) shows that the slope for low-productivity firms is more or less flat

and mostly insignificant, whereas the downward slope for high-productivity firms is steep and significant over most parts of the distribution of the tax differential. The confidence intervals don't overlap again towards the middle of the distribution.

We test the robustness of these results as follows. We first exclude all tax havens used in Davies et al., 2018. These include the Bahamas, Bermuda, Cayman Islands, Cyprus, Hong Kong, Ireland, Luxembourg, Malta, and Singapore. Next, we exclude all tax havens defined in Hines and Rice, 1994. Third, we focus on the 87 parties that are signatory countries of the OECD BEPS agreement as of 12 January 2019. Fourth, we focus on the set of EU countries and exclude all other countries in the sample. The results are reported in Table 5. Overall, they are strikingly similar in magnitude to the results shown in tables 3 and 4, while leading to coefficients that display even higher levels of statistical significance.

5.4 Structural estimation

As a final step, we are interested in the parameters α and β of the concealment cost function, which allows us to perform counter-factual analysis. The theoretical model allows us to identify relationships that pin down these parameters, i.e., equations (15) and (19). We take logs of the aforementioned equations which yields

$$\begin{aligned} log(x_f^U) = & \frac{1}{\beta-1} log\Big(\epsilon(1-\tau_d) - (1-\tau_f)\Big) - \frac{1}{\beta-1} log\,(1-\tau_f) \\ & - \frac{1}{\beta-1} log\,(\alpha) - \frac{\alpha-1}{\beta-1} log(p_A - p_U) + u, \end{aligned} \tag{24}$$

which can be used to identify α and β. Note that we need the deviation from the competitive price, $p_A - p_U$, to estimate the parameters. In the previous section, we have employed extensive fixed effects, which not only account for a variety of factors as described in the previous subsection, but allow us to tease out the difference between intra-firm and ALP import prices and quantities, because average firm-product-year prices and average origin-product-year prices are accounted for by the fixed effects. We can thus construct the price wedge $p_A - p_U$ from the predicted prices resulting from regressions based on equation (22). We use iceberg trade costs from Egger (2014).

6 Conclusion

In this paper we have developed a theoretical model of transfer pricing. Firms that shift profits to low tax destinations face concealment costs and trade costs. The fixed part of the

concealment cost function as well as the trade costs limit the number of firms (or possible location of firms) that are able to use transfer pricing. The specific functional form of the concealment cost function drives the intensive margin, i.e., the elasticity of the price wedge and the quantity of goods shipped. We find that the profits shifted increase with firm productivity, i.e., more productive firms deviate less from the arms' length price, but ship a larger quantity of goods. A greater tax differential increases profits shifting on average in terms of price deviation and quantity. Prices decline on average by 0.4% to 0.7% if the tax wedge increases by one percentage point, while the quantity increases by around 2.4%. Our theoretical model is consistent with empirical findings from reduced form estimations using Swiss transaction-level import data, as well as data about firm-level productivity. Finally, the theoretical model allows to estimate a theory-consistent concealment cost function which can be used for counterfactual analysis. Our findings are consistent with compliance (Jost et al., 2014; Bauer and Langenmayr, 2013; Becker and Davies, 2014; Rathke, 2015) or managerial incentives frameworks (Baldenius et al., 2004; Koethenbuerger and Stimmelmayr, 2015).

References

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer, "Identification properties of recent production function estimators," *Econometrica*, 2015, *83* (6), 2411–2451.

Baldenius, Tim, Nahum D. Melumad, and Stefan Reichelstein, "Integrating managerial and tax objectives in transfer pricing," *The Accounting Review*, 2004, *79* (3), 591–615.

Bauer, Christian J. and Dominika Langenmayr, "Sorting into outsourcing: Are profits taxed at a gorilla's arm's length?," *Journal of International Economics*, 2013, *90* (2), 326–336.

Becker, Johannes and Ronald B. Davies, "A negotiation-based model of tax-induced transfer pricing," CESifo Working Paper Series 4892, CESifo 2014.

Bernard, Andrew B., J. Bradford Jensen, and Peter K. Schott, "Transfer pricing by U.S.-based multinational firms," NBER Working Papers 12493, National Bureau of Economic Research, Inc. 2006.

Clausing, Kimberly A., "Multinational firm tax avoidance and Tax Policy," *National Tax Journal*, 2009, *62* (4), 703–725.

Cristea, Anca D. and Daniel X. Nguyen, "Transfer pricing by multinational firms: New evidence from foreign firm ownerships," *American Economic Journal: Economic Policy*, 2016, *8* (3), 170–202.

Davies, Ronald B, Julien Martin, Mathieu Parenti, and Farid Toubal, "Knocking on Tax Haven's Door: Multinational Firms and Transfer Pricing," *The Review of Economics and Statistics*, 2018, *100* (1), 120–134.

Dharmapala, Dhammika, "What do we know about base erosion and profit shifting? A review of the empirical literature," *Fiscal Studies*, December 2014, *35* (4), 421–448.

_ **and Nadine Riedel**, "Earnings shocks and tax-motivated income-shifting: Evidence from European multinationals," *Journal of Public Economics*, January 2013, *97*, 95–107.

Dyreng, Scott D.D., Michelle Hanlon, and Edward L. Maydew, "Where do firms manage earnings?," *Review of Accounting Studies*, 2012, *17* (3), 649 – 687.

Egger, Peter and Marko Koethenbuerger, "Hosting multinationals: Economic and fiscal implications," *Aussenwirtschaft*, 02 2016, *67* (01), 45–69.

_ **, Christian Keuschnigg, Valeria Merlo, and Georg Wamser**, "Corporate taxes and internal borrowing within multinational firms," *American Economic Journal: Economic Policy*, 2014, *6* (2), 54–93.

_ **, Raymond Riezman, and Benedikt Rydzek**, "Multi-unit firms and their scope and location decision," December 2015.

_ **, Wolfgang Eggert, and Hannes Winner**, "Saving taxes through foreign plant ownership," *Journal of International Economics*, 2010, *81* (1), 99–108.

Gandhi, Amit, Salvador Navarro, and David Rivers, "On the identification of production functions: How heterogeneous is productivity," 2016.

Griffith, R., H. Miller, and M. O'Connell, "Ownership of intellectual property and corporate taxation," *Journal of Public Economics*, April 2014, *112*, 12–23.

Helpman, Elhanan, Marc J. Melitz, and Stephen R. Yeaple, "Export versus FDI with heterogeneous firms," *American Economic Review*, 2004, *94* (1), 300–316.

Hines, James R. Jr. and Eric M. Rice, "Fiscal paradise: foreign tax havens and American business," *Quarterly Journal of Economics*, 1994, *109* (1), 149–182.

Huizinga, H. and L. Laeven, "International profit shifting within multinationals: A multi-country perspective," *Journal of Public Economics*, June 2008, *92* (5–6), 1164–1182.

Jost, Sven P., Michael Pfaffermayr, and Hannes Winner, "Transfer pricing as a tax compliance risk," *Accounting and Business Research*, 2014, *23* (1), 171–197.

Karkinsky, T. and N. Riedel, "Corporate taxation and the choice of patent location within multinational firms," *Journal of International Economics*, September 2012, *88* (1), 176–185.

Keuschnigg, Christian and Michael P. Devereux, "The arm's length principle and distortions to multinational firm organization," *Journal of International Economics*, 2013, *89* (2), 432 – 440.

Koethenbuerger, Marko and Michael Stimmelmayr, "Tax-induced transfer pricing and corporate agency costs," 2015. mimeo.

Krugman, Paul, "Scale Economies, Product Differentiation, and the Pattern of Trade," *The American Economic Review*, 1980, *70* (5), 950–959.

Liu, Li, Tim Schmidt-Eisenlohr, and Dongxian Guo, "International transfer pricing and tax avoidance: Evidence from linked trade-tax statistics in the UK," CESifo Working Paper Series 6594, CESifo 2017.

Melitz, Marc J., "The impact of trade on intra-industry reallocations and aggregate industry productivity," *Econometrica*, November 2003, *71* (6), 1695–1725.

_ **and Gianmarco I.P. Ottaviano**, "Market size, trade, and productivity," *The Review of Economic Studies*, 2008, *75* (1), 295–316.

OECD, "Review of comparability and of profit methods: Revision of Chapters I-III of the transfer pricing guidelines," Technical Report, OECD July 2010.

Petrin, A., B.P. Poi, and J. Levinsohn, "Production function estimation in Stata using inputs to control for unobservables," *Stata Journal*, 2004, *4*, 113–123.

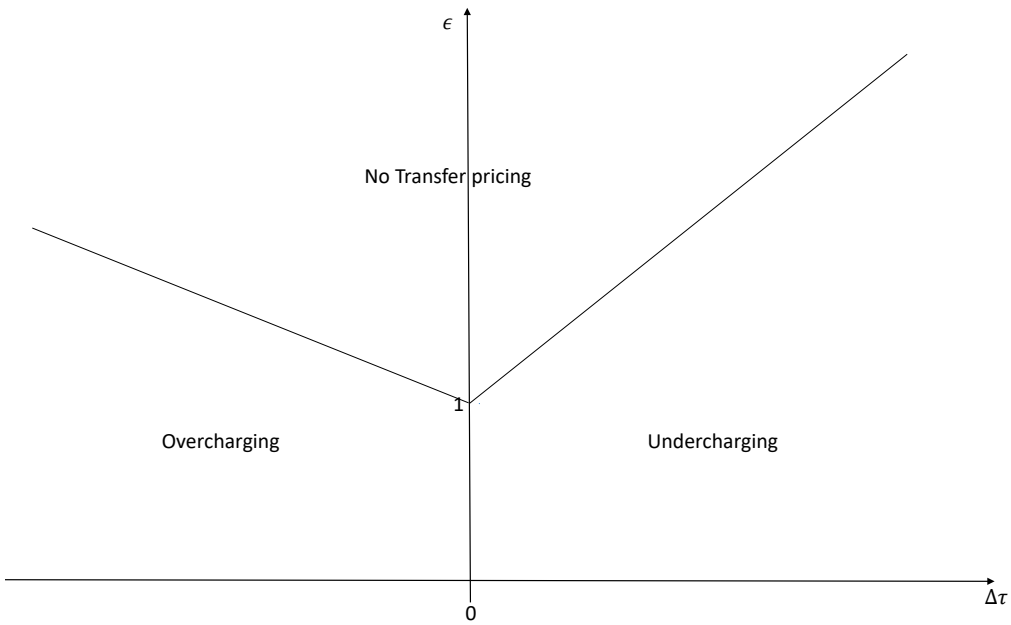
Rathke, Alex A. T., "Transfer pricing manipulation, tax penalty cost and the impact of foreign profit taxation," EconStor Conference Papers 129075, ZBW - Leibniz Information centre for Economics 2015.

Schindler, D. and G. Schjelderup, "Multinationals and income shifting by debt," *International Journal of the Economics of Business*, 2016, *23* (3), 263–286.

Torslov, Thomas R., Ludvig S. Wier, and Gabriel Zucman, "The Missing Profits of Nations," Working Paper 24701, National Bureau of Economic Research June 2018.

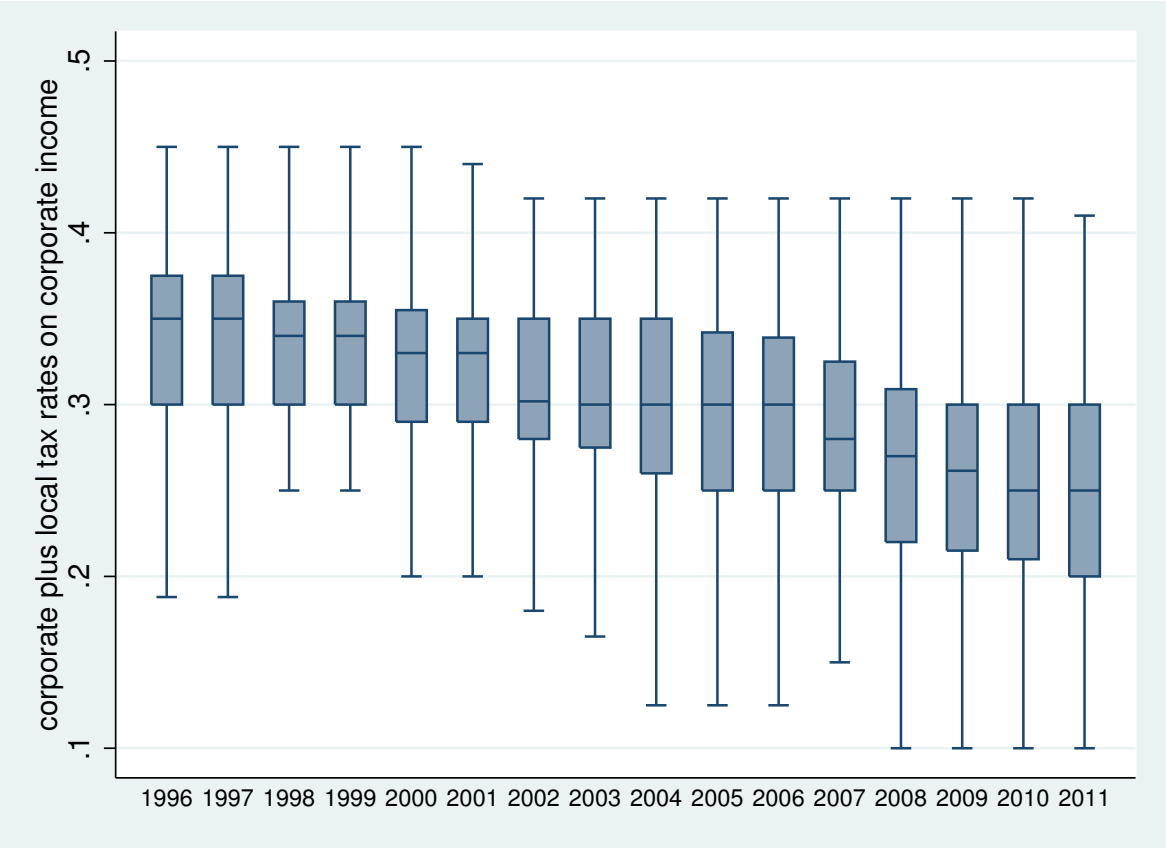
Tables and figures

Figure 1: Condition for ϵ to ensure transfer pricing.



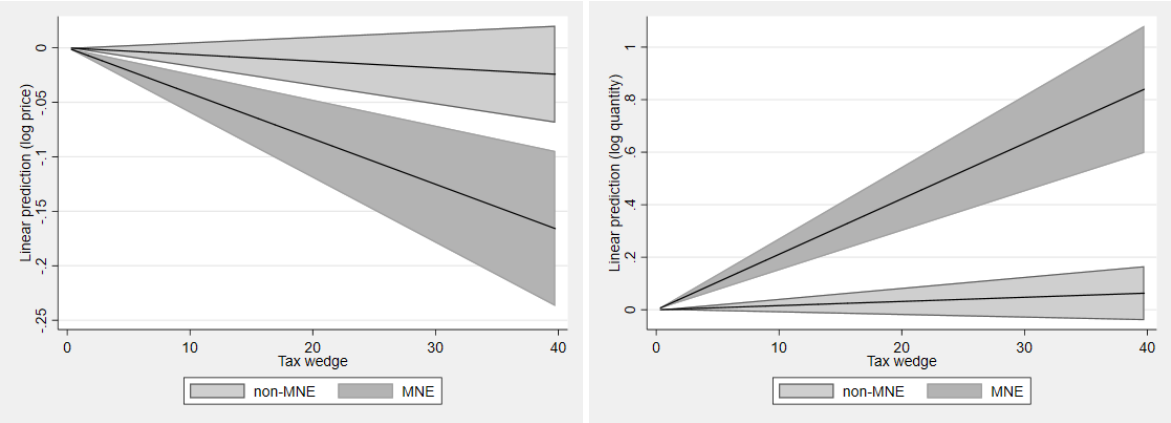
Note: We fix $\tau_d = 0.5$ and only vary τ_f. We assume that $\alpha = 0.75$ and $\beta = 0.25$. For ϵ above the line, transfer pricing is not feasible for firms, independent of their productivity.

Figure 2: Corporate profit tax. 79 countries, 1996 - 2011



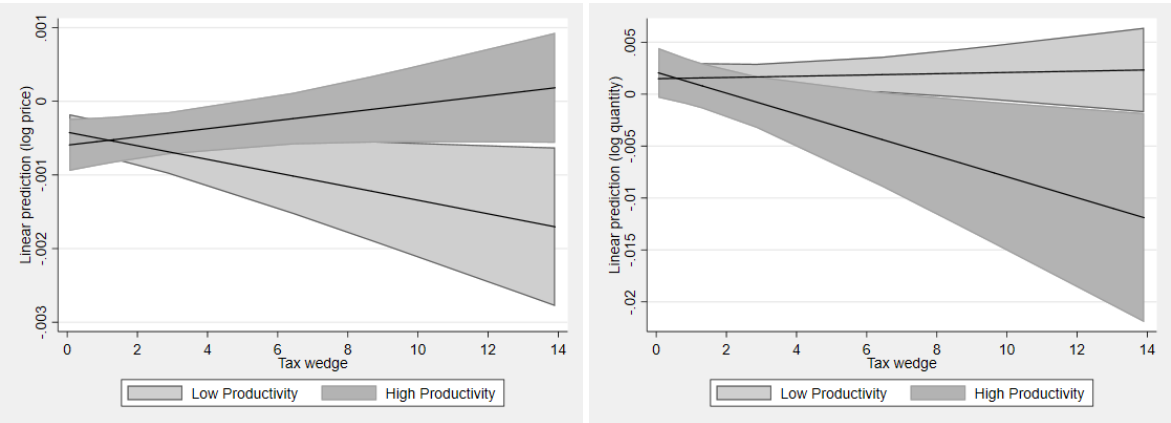
Note: For corporate profit taxes, we utilize the maximum tax rate levied at the national level on corporate profit in a country of residence. In federal states, the total corporate tax rate is calculated as the weighted average of the local (sub-national) taxes combined with federal tax rates (e.g., for Germany or Canada as reported by the OECD) or the tax rate prevailing in the economic center (e.g. for Switzerland, where the rates of the canton of Zurich are taken). The primary sources for corporate profit tax rates are the following: Ernest and Young *Worldwide Corporate Tax Guide* 1998-2012; Coopers and Lybrand *International Tax Summaries* 1996-1997; International Bureau of Fiscal Documentation *Global Corporate Tax Handbook* 2007-2012; Price Waterhouse Coopers *Corporate Taxes - Worldwide Summaries* 1999-2000, 2001-2003, 2012-2013; OECD *www.taxfoundation.org*.

Figure 3: Effect of tax for intra-firm versus arm's length transactions on log prices and log quantities of import transactions



Notes: The figures plot the effect of taxes (for different percentiles of the distribution of the tax differential) for intra-firm versus arm's length transactions.

Figure 4: Effect of tax for high- versus low-productivity firms on log prices and log quantities of import transactions



Notes: The figures plot the effect of taxes firms with high (mean plus one standard deviation) versus low (mean minus one standard deviation) firms.

Table 1: Corporate tax rates in Switzerland, 2015

Lucerne	13.63	St. Gallen	17.40
Nidwalden	12.73	Aargau	17.25
Obwalden	14.16	Fribourg	21.86
Appenzell A.	13.55	Ticino	22.55
Appenzell I.	14.16	Basle-Land	20.73
Zug	13.87	Jura	23.31
Schwyz	14.86	Zurich	21.15
Uri	15.12	Valais	19.76
Glarus	17.44	Berne	20.49
Schaffhausen	17.40	Solothurn	19.64
Thurgau	16.43	Basle	24.79
Grisons	19.94	Vaud	22.79
Neuchâtel	17.31	Geneva	24.67

Notes: Taxes (cantonal, municipality, church and federal taxes) in % of net profits, average over all yield classes, for incorporated companies with capital and reserves of 2 million Swiss Francs. The tax burden is calculated for the cantonal capital. Source: ESTV.

Table 2: Summary statistics

	Mean	SD	Min	Max
A. Statistics on taxes				
Pairs *cft* with $I_{\Delta\tau_{fct}>0}$				21,773
Pairs *cft* with $I_{\Delta\tau_{fct}<0}$				8,777
$\Delta\tau_{fct}$	4.159	9.782	-26.715	39.750
$\Delta\tau_{fct}$ ($I_{\Delta\tau_{fct}>0}$)	9.081	5.562	0.026	39.750
$\Delta\tau_{fct}$ ($I_{\Delta\tau_{fct}<0}$)	-8.050	6.871	-26.715	-0.001
B. Statistics on import transactions				
No. Firms				397,485
No. Cantons				26
No. Products				8,401
No. Exporters				1,789,852
No. Origins				189
No. Years				10
MNE_{ijt}				134,470
Non-MNE_{ijt}				20,587,479
Log price	3.491	2.120	-14.000	23.556
Log quantity	3.553	3.193	-6.908	20.572
Log price for MNE_{ijt} ($I_{\Delta\tau_{fct}>0}$)	3.588	1.963	-7.544	20.129
Log price for non-MNE_{ijt} ($I_{\Delta\tau_{fct}>0}$)	3.458	2.129	-14.000	23.556
Log quantity for MNE_{ijt} ($I_{\Delta\tau_{fct}>0}$)	4.137	3.038	-6.908	17.060
Log quantity for non-MNE_{ijt} ($I_{\Delta\tau_{fct}>0}$)	3.672	3.205	-6.908	20.572
Log price for MNE_{ijt} ($I_{\Delta\tau_{fct}<0}$)	3.645	1.964	-9.944	19.909
Log price for non-MNE_{ijt} ($I_{\Delta\tau_{fct}<0}$)	3.627	2.079	-12.192	23.085
Log quantity for MNE_{ijt} ($I_{\Delta\tau_{fct}<0}$)	3.764	2.998	-6.908	17.151
Log quantity for non-MNE_{ijt} ($I_{\Delta\tau_{fct}<0}$)	3.033	3.094	-6.908	19.746
Observations				20,721,949
C. Firm-level statistics				
Log productivity (ϕ_{it})	7.220	1.434	-0.308	14.922
Observations with ϕ_{it}				2,251,885

Notes: Tax statistics stem from the Federal Tax Administration *ESTV*, and Egger et al., 2015. The import summary statistics correspond to data pooled by firm (i), canton (c), product (k), exporter firm (j), origin (f) and year (t) over the period 2006-2015 for import transactions obtained from the Swiss Federal Customs Administration (*EZV*, Oberzolldirektion). Log price refers to the unit value, calculated as the log of volume divided by quantity on data containing observations about single transactions. The data correspond to a match of trade transactions data with Amadeus firm-level data from Bureau van Dijk. Log ϕ_{it} is calculated according to Petrin et al. (2004).

Table 3: Effect of taxes on prices and quantities of intra-firm versus arm's length import transactions

	(1)	(2)	(3)	(4)	(5)
A. Log price					
$I_{\Delta\tau_{fct}>0} \times \Delta\tau_{fct}$		0.002	0.008*	0.003	0.001
		(0.002)	(0.005)	(0.004)	(0.002)
$I_{\Delta\tau_{fct}>0} \times MNE_{ijt} \times \Delta\tau_{fct}$	-0.004***	-0.004***	-0.004***	-0.004***	-0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$I_{\Delta\tau_{fct}<0} \times \Delta\tau_{fct}$		0.003*	0.012**	0.004	0.003
		(0.002)	(0.005)	(0.004)	(0.002)
$I_{\Delta\tau_{fct}<0} \times MNE_{ijt} \times \Delta\tau_{fct}$	0.008***	0.008***	0.008***	0.008***	0.009***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Obs.	20,721,949	20,721,949	6,598,142	20,626,913	13,855,001
Fixed effects *ifk*	3,167,558	3,167,558	783,917	3,168,404	2,269,681
Fixed effects *fkt*	462,628	462,628	296,559	461,886	342,468
Fixed effects *ikt*	5182979	5182979	1321395	5154631	3765725
No. cl.	23,165	23,165	18,708	23,100	17,809
R^2	0.839	0.839	0.814	0.840	0.869
B. Log quantity					
$I_{\Delta\tau_{fct}>0} \times \Delta\tau_{fct}$		-0.001	-0.016	0.009	-0.005
		(0.005)	(0.014)	(0.010)	(0.007)
$I_{\Delta\tau_{fct}>0} \times MNE_{ijt} \times \Delta\tau_{fct}$	0.024***	0.024***	0.024***	0.024***	0.024***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
$I_{\Delta\tau_{fct}<0} \times \Delta\tau_{fct}$		-0.004	-0.017	0.007	-0.010
		(0.005)	(0.013)	(0.010)	(0.007)
$I_{\Delta\tau_{fct}<0} \times MNE_{ijt} \times \Delta\tau_{fct}$	-0.055***	-0.055***	-0.056***	-0.056***	-0.053***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.010)
Obs.	20,721,949	20,721,949	6,598,142	20,626,913	13,855,001
Fixed effects *ifk*	3,167,558	3,167,558	783,917	3,168,404	2,269,681
Fixed effects *fkt*	462,628	462,628	296,559	461,886	342,468
Fixed effects *ikt*	5182979	5182979	1321395	5154631	3765725
No. cl.	23,165	23,165	18,708	23,100	17,809
R^2	0.817	0.817	0.777	0.818	0.779

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively. All equations are estimated by pooled OLS with absorbed fixed effects (*ifk* importer-origin-product, *fkt* origin-product-year, *ikt* importer-product-year), and standard errors clustered at the level of canton-origin-year. Dependent variables: log unit value per import transaction (log price) and log quantity per import transaction (log quantity), based on data aggregated at the level of firm-canton-product-exporter-origin-year. Regressions in (1) include the interaction term only; (2) estimates eq. (22); (3) estimates eq. (22), based on firms that have at least one MNE transaction; (4) employs *icfk* (importer-canton-origin-product) instead of *ifk* (importer-origin-product) fixed effects; (5) reports results based on *EZV* product numbers (*Tarifnummern*) 39-92, 94, 95 and 96.

Table 4: Effect of taxes and productivity on prices and quantities of intra-firm versus arm's length import transactions

	(1)	(2)	(3)	(4)	(5)	(6)
A. Log price						
$\text{MNE}_{\text{ijt}} \times \phi_{\text{it}}$	-0.005	-0.005	-0.006	-0.008***	-0.005	-0.007
	(0.005)	(0.005)	(0.005)	(0.001)	(0.003)	(0.005)
$\text{I}_{\Delta\tau_{\text{fct}}>0} \times \text{MNE}_{\text{ijt}} \times \Delta\tau_{\text{fct}}$	-0.022***	-0.023***	-0.022***	-0.022***		-0.021***
	(0.007)	(0.007)	(0.007)	(0.005)		(0.008)
$\text{I}_{\Delta\tau_{\text{fct}}>0} \times \text{MNE}_{\text{ijt}} \times \Delta\tau_{\text{fct}} \times \phi_{\text{it}}$	0.003**	0.003***	0.003**	0.003***		0.002**
	(0.001)	(0.001)	(0.001)	(0.001)		(0.001)
$\text{I}_{\Delta\tau_{\text{fct}}<0} \times \text{MNE}_{\text{ijt}} \times \Delta\tau_{\text{fct}}$	0.022	0.024	0.014		0.023	0.025
	(0.068)	(0.070)	(0.069)		(0.057)	(0.098)
$\text{I}_{\Delta\tau_{\text{fct}}<0} \times \text{MNE}_{\text{ijt}} \times \Delta\tau_{\text{fct}} \times \phi_{\text{it}}$	-0.004	-0.004	-0.003		-0.004	-0.001
	(0.009)	(0.010)	(0.009)		(0.007)	(0.013)
Obs.	2,657,561	1,080,056	2,650,654	2,657,561	2,657,561	1,994,067
Fixed effects *ifk*	551,650	192,562	551,429	551,650	551,650	440,423
Fixed effects *fkt*	135,442	83,297	135,272	135,442	135,442	107,199
No. cl.	8,040	5,863	7,978			6,041
R^2	0.821	0.794	0.821	0.821	0.821	0.837
B. Log quantity						
$\text{MNE}_{\text{ijt}} \times \phi_{\text{it}}$	0.061**	0.060**	0.062**	0.005**	0.059***	0.067**
	(0.028)	(0.028)	(0.028)	(0.002)	(0.005)	(0.027)
$\text{I}_{\Delta\tau_{\text{fct}}>0} \times \text{MNE}_{\text{ijt}} \times \Delta\tau_{\text{fct}}$	0.132***	0.135***	0.132***	0.132***		0.132***
	(0.047)	(0.048)	(0.047)	(0.008)		(0.051)
$\text{I}_{\Delta\tau_{\text{fct}}>0} \times \text{MNE}_{\text{ijt}} \times \Delta\tau_{\text{fct}} \times \phi_{\text{it}}$	-0.022***	-0.022***	-0.022***	-0.022***		-0.023***
	(0.008)	(0.008)	(0.008)	(0.001)		(0.008)
$\text{I}_{\Delta\tau_{\text{fct}}<0} \times \text{MNE}_{\text{ijt}} \times \Delta\tau_{\text{fct}}$	-0.152	-0.160	-0.149		-0.145	-0.195
	(0.122)	(0.121)	(0.123)		(0.095)	(0.215)
$\text{I}_{\Delta\tau_{\text{fct}}<0} \times \text{MNE}_{\text{ijt}} \times \Delta\tau_{\text{fct}} \times \phi_{\text{it}}$	0.022	0.023	0.022		0.016	0.026
	(0.016)	(0.016)	(0.016)		(0.012)	(0.026)
Obs.	2,657,561	1,080,056	2,650,654	2,657,561	2,657,561	1,994,067
Fixed effects *ifk*	551,650	192,562	551,429	551,650	551,650	440,423
Fixed effects *fkt*	135,442	83,297	135,272	135,442	135,442	107,199
No. cl.	8,040	5,863	7,978			6,041
R^2	0.792	0.758	0.792	0.792	0.792	0.755

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively. All equations are estimated by pooled OLS with absorbed fixed effects (*ifk* importer-origin-product, *fkt* origin-product-year), and standard errors clustered at the level of canton-origin-year. Dependent variables: log unit value per import transaction (log price) and log quantity per import transaction (log quantity), based on data aggregated at the level of firm-canton-product-exporter-origin-year. Regressions in (1) estimates eq. (23); (2) estimates eq. (23), based on firms that have at least one MNE transaction; (3) employs *icfk* (importer-canton-origin-product) instead of *ifk* (importer-origin-product) fixed effects; (4) focuses on $\Delta\tau_{\text{fct}} > 0$; (5) focuses on $\Delta\tau_{\text{fct}} < 0$; and (6) reports results based on *EZV* product numbers (*Tarifnummern*) 39-92, 94, 95 and 96.

Table 5: Robustness analysis of the effect of taxes and productivity on prices and quantities of intra-firm versus arm's length import transactions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Log price								
$I_{\Delta\tau_{fct}<0} \times \Delta\tau_{fct}$	0.003*	0.003*	0.004**	0.002				
	(0.002)	(0.002)	(0.002)	(0.002)				
$I_{\Delta\tau_{fct}>0} \times \Delta\tau_{fct}$	0.002	0.002	0.002	0.001				
	(0.002)	(0.002)	(0.002)	(0.002)				
$I_{\Delta\tau_{fct}<0} \times MNE_{ijt} \times \Delta\tau_{fct}$	0.008***	0.009***	0.008***	0.010***				
	(0.002)	(0.002)	(0.002)	(0.002)				
$I_{\Delta\tau_{fct}>0} \times MNE_{ijt} \times \Delta\tau_{fct}$	-0.004***	-0.004***	-0.004***	-0.004***				
	(0.001)	(0.001)	(0.001)	(0.001)				
$MNE_{ijt} \times \phi_{it}$					-0.005	-0.005	-0.008*	-0.008
					(0.005)	(0.005)	(0.005)	(0.005)
$I_{\Delta\tau_{fdt}>0} \times MNE_{ijt} \times \Delta\tau_{fdt}$					-0.022***	-0.022***	-0.029***	-0.029***
					(0.007)	(0.007)	(0.008)	(0.008)
$I_{\Delta\tau_{fdt}<0} \times MNE_{ijt} \times \Delta\tau_{fdt}$					0.088*	0.086	0.012	0.150
					(0.051)	(0.053)	(0.071)	(0.112)
$I_{\Delta\tau_{fdt}>0} \times MNE_{ijt} \times \Delta\tau_{fdt} \times \phi_{it}$					0.003**	0.003**	0.004***	0.004***
					(0.001)	(0.001)	(0.001)	(0.001)
$I_{\Delta\tau_{fdt}<0} \times MNE_{ijt} \times \Delta\tau_{fdt} \times \phi_{it}$					-0.015**	-0.015**	-0.003	-0.019
					(0.007)	(0.007)	(0.010)	(0.016)
Obs.	20,356,616	20,348,845	19,523,418	17,980,279	2,619,117	2,618,732	2,537,568	2,375,701
Fixed effects *ifk*	3,093,649	3,091,626	2,966,098	2,704,327	542,633	542,500	525,232	491,437
Fixed effects *fkt*	447,905	446,851	418,161	336,022	132,145	132,035	125,811	108,920
No. cl.	21,604	20,577	15,402	6,804	7,433	7,315	6,518	3,584
R^2	0.838	0.838	0.839	0.834	0.820	0.820	0.820	0.817
B. Log quantity								
$I_{\Delta\tau_{fct}<0} \times \Delta\tau_{fct}$	-0.005	-0.005	-0.006	-0.007				
	(0.005)	(0.005)	(0.005)	(0.005)				
$I_{\Delta\tau_{fct}>0} \times \Delta\tau_{fct}$	-0.002	-0.002	-0.002	-0.004				
	(0.005)	(0.005)	(0.005)	(0.005)				
$I_{\Delta\tau_{fct}<0} \times MNE_{ijt} \times \Delta\tau_{fct}$	-0.054***	-0.055***	-0.056***	-0.059***				
	(0.009)	(0.009)	(0.008)	(0.009)				
$I_{\Delta\tau_{fct}>0} \times MNE_{ijt} \times \Delta\tau_{fct}$	0.024***	0.024***	0.020***	0.020***				
	(0.004)	(0.004)	(0.005)	(0.005)				
$MNE_{ijt} \times \phi_{it}$					0.061**	0.061**	0.095***	0.105***
					(0.028)	(0.028)	(0.032)	(0.033)
$I\Delta\tau_{fdt} > 0 \times MNE_{ijt} \times \Delta\tau_{fdt}$					0.132***	0.132***	0.156***	0.159***
					(0.047)	(0.047)	(0.053)	(0.054)
$I\Delta\tau_{fdt} < 0 \times MNE_{ijt} \times \Delta\tau_{fdt}$					-0.253**	-0.266**	-0.119	-0.115
					(0.112)	(0.117)	(0.113)	(0.175)
$I\Delta\tau_{fdt} > 0 \times MNE_{ijt} \times \Delta\tau_{fdt} \times \phi_{it}$					-0.022***	-0.022***	-0.029***	-0.030***
					(0.008)	(0.008)	(0.010)	(0.010)
$I\Delta\tau_{fdt} < 0 \times MNE_{ijt} \times \Delta\tau_{fdt} \times \phi_{it}$					0.037***	0.038***	0.023	0.030
					(0.014)	(0.015)	(0.015)	(0.018)
Obs.	20,356,616	20,348,845	19,523,418	17,980,279	2,619,117	2,618,732	2,537,568	2,375,701
Fixed effects *ifk*	3,093,649	3,091,626	2,966,098	2,704,327	542,633	542,500	525,232	491,437
Fixed effects *fkt*	447,905	446,851	418,161	336,022	132,145	132,035	125,811	108,920
No. cl.	21,604	20,577	15,402	6,804	7,433	7,315	6,518	3,584
R^2	0.817	0.817	0.816	0.812	0.791	0.791	0.790	0.786

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively. All equations are estimated by pooled OLS with absorbed fixed effects (*ift*, *ifk* importer-origin-product, *fkt* origin-product-year in Columns 1–4, and *ifk* importer-origin-product, *fkt* origin-product-year in Columns 5–8), and standard errors clustered at the level of canton-origin-year. Columns 1–4 exclude productivity, and columns 5–8 include productivity. Dependent variables: log unit value per import transaction (log price) and log quantity per import transaction (log quantity), based on data aggregated at the level of firm-canton-product-exporter-origin-year. Regressions in (1) and (5), excluding the tax havens as in Davies et al., 2018; (2) and (6) exclude the tax havens as in Hines and Rice, 1994; (3) and (7) focus on countries which adhere to the OECD BEPS framework; and (4) and (8) are based on EU countries only.

A Expected fines of tax evasion

In this section we derive the proposed concealment cost functions in equations (1) and (2) from the optimal behavior of a tax authority that maximizes expected tax revenues. Without loss of generality we focus only on the undercharging case. Deriving the concealment cost function for the overcharging case follows analogously. The tax authority decides how many revenue officers it wants to employ to audit a parent firm given some ex-ante observable

variables. The detection probability of illegal transfer pricing increases with the price wedge between the competitive arm's length price and the undercharging price. Both taken as given from the perspective of the tax authority. Moreover, the probability increases (for a given price wedge) with the number of revenue officers are employed at the audit of the firm. We assume that the tax authority randomly audit firms, but a firm is selected for an audit the tax authority can costlessly observe the price wedge and the traded quantity. Still the tax authority has to confirm that the firm indeed applies illegal transfer pricing. Thus, the expected tax revenues, E_T, from an audit are

$$E_T = \left(\frac{L(p_A - p_U)^{\varsigma}}{L(p_A - p_U)^{\varsigma} + 1}\right)(\tau_f(p_A - p_U)x_f\varphi) - wL, \tag{25}$$

where the first parenthesis gives the detection probability as a function of revenue officers employed, the price wedge, and an elasticity parameter $\varsigma > 0$. The second parenthesis corresponds to the evaded tax revenues multiplied by a fine markup, $\varphi > 1$, and wL are the wage costs of revenue officers.

The optimal amount of tax officers is given by

$$L(p_A - p_U)^{\varsigma} + 1 = \left(\frac{\tau_f x\varphi(p_A - p_U)^{\varsigma+1}}{w}\right)^{\frac{1}{2}}. \tag{26}$$

Assume that the probability of a firm to get audited is $\frac{\overline{L}}{ML}$, where $\overline{L}$ is the (inelastic) total number of tax officers at the tax authority and M is the total number of firms in the economy that the tax authority could audit. Then the expected fine from the perspective can be written as

$$E_\Pi = \frac{\overline{L}}{ML}\frac{L(p_A - p_U)^{\varsigma}}{L(p_A - p_U)^{\varsigma} + 1}\tau_f(p_A - p_U)x_f\varphi = \frac{\overline{L}}{M}(x\varphi\tau_f)^{\frac{1}{2}}(p_A - p_U)^{\frac{\varsigma+1}{2}}, \tag{27}$$

where we substituted the optimal number of tax officers from equation (26). This is equivalent to the concealment costs function in equation (1), scaled by $\frac{\overline{L}}{M}(\varphi\tau_f)^{\frac{1}{2}}$, and $\beta = \frac{\varsigma+1}{2}$, and $\alpha = \frac{1}{2}$.

B Overcharging MNE

If the tax rate in the domestic country is higher than in the foreign country $\tau_d > \tau_f$, firms have incentives to shift profits to the foreign country and thus the foreign affiliate overcharges the domestic firm, $p_O > 1/\phi$. Total net profits are

$$\begin{aligned}\pi^O(\omega) =&(1-\tau_d)\left(p(\omega)y(\omega)-\frac{x_d}{\phi}-p_O\epsilon x_f\right)+(1-\tau_f)\left(p_O-\frac{1}{\phi}\right)x_f\\ &-(1-\tau_d)(p_O-p_A)^\alpha x_f^\beta)-F_O-F.\end{aligned} \tag{28}$$

The first order conditions of the firm's maximization problem are:

$$\begin{aligned}\frac{\partial\pi^O(\omega)}{\partial x_d} =&E-2(x_d+x_f)-\frac{1}{\phi}=0,\\ \frac{\partial\pi^O(\omega)}{\partial x_f} =&(1-\tau_d)(E-2(x_d+x_f)-p_O\epsilon)+(1-\tau_f)\left(p_O-\frac{1}{\phi}\right)-(1-\tau_d)\beta x_f^{\beta-1}(p_O-p_A)^\alpha=0,\\ \frac{\partial\pi^O(\omega)}{\partial p_O} =&-(1-\tau_d)\epsilon x_f+(1-\tau_f)x_f-(1-\tau_d)\alpha(p_O-p_A)^{\alpha-1}x_f^\beta=0\end{aligned}$$

Solving for the optimal p_O yields

$$p_O=\frac{\beta}{\beta-\alpha}\left(p_A+\frac{\alpha}{\beta}\frac{\tau_d-\tau_f}{\phi\zeta}\right). \tag{29}$$

Note that the overcharging price is higher than the arms' length price if $\zeta\leq 0$. This is the case if

$$\epsilon<\frac{1-\tau_f}{1-\tau_d}, \tag{30}$$

where the RHS is greater than one as $\tau_d>\tau_f$. Similarly to the undercharging case transportation costs limit transfer pricing if the tax differential is too small.

Similar to the undercharging case we derive the comparative statics for overcharging:

$$\partial p_O/\partial\phi<0,\qquad \partial p_O/\partial\epsilon<0,$$

$$\partial p_O/\partial\tau_f<0,\qquad \partial p_O/\partial\tau_d>0,\qquad \partial^2 p_U/\partial\tau_f\partial\phi>0,$$

where we use that for $p_A=\frac{1}{\phi}$, $\alpha>\beta$ and $\zeta<0$ to ensure that $p_O>\frac{1}{\phi}$. In the overcharging case high-productivity firms increase the transfer price less than low-productivity firms.

Solving for x_f yields

$$x_f^O=\left(\frac{(1-\tau_f)-\epsilon(1-\tau_d)}{(1-\tau_d)\alpha(p_O-p_A)^{\alpha-1}}\right)^{\frac{1}{\beta-1}}. \tag{31}$$

Again substituting p_O from equation (17) yields the equilibrium traded quantity:

$$x_f^O=\left(\frac{-\zeta}{(1-\tau_f)\alpha\left(\frac{\alpha}{\alpha-\beta}\left(p_A+\frac{\tau_d-\tau_f}{\phi\zeta)}\right)\right)^{\alpha-1}}\right)^{\frac{1}{\beta-1}}. \tag{32}$$

Again, bigger firms trade more goods at a smaller price differential relative to the competitive price. Moreover, $\partial x_f^O/\partial \tau_d > 0$ and $\partial x_f^O/\partial \tau_f < 0$, which implies that a higher tax difference leads to increased imports of the over-priced foreign inputs. The cutoff productivity can be determined analogously to the undercharging case.

C Proofs

C.1 Undercharging price condition

The undercharging price as defined by equation (13) will always satisfy the following inequalities:

$$0 \leq p_U \leq p_A,$$

if $p_A = \frac{1}{\phi}$, trade costs $\epsilon > 1$, $\alpha > \beta$ and $\zeta > 0$ As by definition of the undercharging case $\tau_d < \tau_f$ it is easy to show that $\zeta > 0$. The undercharging price is lower than the ALP if

$$p_U - p_A < 0,$$

$$\frac{\alpha}{\beta - \alpha\left(\frac{\beta}{\alpha}p_A - \frac{\tau_f - \tau_d}{\phi\zeta}\right) - p_A < 0},$$

where we use that $p_A = \frac{1}{\phi}$. Simplifying yields

$$\frac{\alpha}{\beta - \alpha}\left(1 - \frac{\tau_f - \tau_d}{\zeta}\right) > 0,$$

using that $\alpha > \beta$ yields

$$\zeta > \tau_f - \tau_d,$$

after substituting ζ and simplifying this gives

$$1 > \tau_d,$$

which holds by definition of $\tau_d \in [0, 1]$.

C.2 Comparative statics

Equation (13) defines the undercharging price. The price deviation is given by

$$p_A - p_U = \frac{1}{\phi} - \frac{\beta}{\beta - \alpha}\frac{1}{\phi} + \frac{\tau_f - \tau_d}{\zeta\phi}\frac{\alpha}{\beta - \alpha}.$$

Price derivatives:

$$\frac{\partial p_A - p_U}{\partial \phi} < 0.$$

$$-\frac{1}{\phi^2} + \frac{\beta}{\beta - \alpha}\frac{1}{\phi^2} - \frac{\tau_f - \tau_d}{\zeta\phi^2}\frac{\alpha}{\beta - \alpha} < 0.$$

Given that $\alpha > \beta$ we have that

$$-(\beta - \alpha) + \beta - \frac{\tau_f - \tau_d}{\zeta}\alpha > 0.$$

Simplifying yields

$$\tau_f - \tau_d > -\zeta,$$

which always holds as $\zeta > 0$ and $\tau_d < \tau_f$ in the undercharging case. More productive firms deviate less from the ALP.

The cross-derivative of the price differential with respect to productivity and the tax rate differential is given by

$$\frac{\partial^2 p_A - p_U}{\partial \phi \partial \tau_d} = \frac{1}{\phi^2}\frac{\alpha}{\beta - \alpha}\left(-\frac{\partial \frac{\tau_f - \tau_d}{\zeta}}{\partial \tau_d}\right).$$

It is easy to show that

$$-\frac{\partial \frac{\tau_f - \tau_d}{\zeta}}{\partial \tau_d} = \frac{-\zeta + (\tau_f - \tau_d)\epsilon}{\zeta^2},$$

which after some simplification yields

$$\frac{(1-\epsilon)(1-\tau_d)}{\zeta^2} < 0.$$

As $\alpha > \beta$ the cross derivative is negative, i.e.,

$$\frac{\partial^2 p_A - p_U}{\partial \phi \partial \tau_d} < 0,$$

as the tax differential decreases the price deviation increases with productivity.

The price deviation with respect to trade costs increase, i.e.,

$$\frac{\partial p_A - p_U}{\partial \epsilon} > 0,$$

as

$$\frac{\alpha}{\beta-\alpha}\frac{\tau_f-\tau_d}{\phi}\left(-\frac{1-\tau_d}{\zeta^2}\right)<0,$$

and $\alpha > \beta$.

Quantity derivatives: To derive the relationship between the undercharging optimal quantity and firm productivity we use equation (15). From above we know that

$$\frac{\partial p_A - p_U}{\partial \phi} < 0.$$

If $1 < \beta < \alpha$ than it is straightforward to see that

$$\frac{\partial x_f^U}{\partial \phi} > 0,$$

and more productive firm use larger quantities to shift profits.

$$\frac{\partial X_f^U}{\partial \tau_d} = \frac{1}{\beta-1}(x_f^U)^{\frac{1}{\beta-1}-1}\left[-((1-\tau_f)\alpha(p_A-p_U)^{\alpha-1}) - (\frac{(1-\tau_d)(1-\epsilon)}{(1-\tau_d)\epsilon-(1-\tau_f)})\alpha(\alpha-1)(p_A-p_U)^{\alpha-2}\frac{\partial(p_A-p_U}{\partial \tau_d}\right.$$

$$x_f^U = \left(\frac{(1-\tau_d)\epsilon-(1-\tau_f)}{(1-\tau_f)\alpha(p_A-p_U)^{\alpha-1}}\right)^{\frac{1}{\beta-1}}. \tag{33}$$

Substituting the optimal undercharging price, p_U, from equation (13) yields

$$x_f^U = \left(\frac{(1-\tau_d)\epsilon-(1-\tau_f)}{(1-\tau_f)\alpha\left(\frac{\alpha}{\alpha-\beta}\left(p_A-\frac{\tau_f-\tau_d}{\phi\zeta}\right)\right)^{\alpha-1}}\right)^{\frac{1}{\beta-1}}. \tag{34}$$

D Data appendix

Table 6: Capital tax burden on holdings in cantonal capitals in Switzerland, 2015 in CHF

Lucerne	500	St. Gallen	1005
Nidwalden	500	Aargau	845
Obwalden	500	Fribourg	651
Appenzell A.	900	Ticino	975
Appenzell I.	500	Basle-Land	308
Zug	374	Jura	630
Schwyz	398	Zurich	687
Uri	500	Valais	400
Glarus	500	Berne	728
Schaffhausen	210	Solothurn	656
Thurgau	837	Basle	1000
Grisons	300	Vaud	3503
Neuchâtel	20	Geneva	1336

Notes: Capital tax burden for holdings with capital and reserves of 2 million Swiss Francs with net profit of 160 thousand Francs. The tax burden is calculated for the cantonal capital. Source: ESTV.