

Tax Evasion and Productivity

Hans Martinez*

Department of Economics, University of Western Ontario

September 3, 2025

Abstract

I propose a novel method to estimate corporate tax evasion through cost overreporting using production functions. I first show how that cost overreporting leads to biased estimates of productivity and the production function parameters. Then, I demonstrate how to recover unbiased estimates of both in the presence of tax evasion. With these estimates on hand, the production function can be inverted to recover the true inputs. Cost overreporting can be estimated as the difference between the observed and the true inputs. I apply this method to a well-known dataset where I find evidence suggesting that cost overreporting is widespread and quantitatively large. My results indicate that ignoring cost overreporting leads to consistently larger elasticities of intermediate inputs. The bias on the intermediate inputs spreads to the elasticities of labor and capital, whose bias direction varies by industry. Finally, I find significant differences in the productivity distributions.

Keywords: Tax Evasion, Cost Overreporting, Production Function Estimation, Productivity

*email: hmarti33@uwo.ca.

Introduction

I propose a novel method to estimate corporate tax evasion through cost overreporting using production functions. I first show how that cost overreporting lead to biased estimates of productivity and production function parameters. Then, I show how to recover unbiased estimates of the production function and productivity in the presence of tax evasion. With the production function parameters and productivity on hand, the production function can be inverted to recover the true inputs. Then, I estimate cost overreporting as the difference between observed and true inputs. I apply this method to a well-known dataset where I find evidence suggesting that cost overreporting is widespread and quantitatively large. I also find that ignoring cost overreporting leads to consistently larger elasticities of intermediate inputs. The bias on the intermediate inputs spreads to the elasticities of labor and capital, whose bias direction varies by industry.

Corporate tax evasion through cost overreporting spreads globally causing governments significant revenue losses. Cost overreporting, however, has been largely overlooked by the literature except for a few recent studies. The literature has mostly relied on manually detecting evaders using administrative data or on experiments where researchers randomly send letters and observing how firms adjust their tax declarations. My approach is complementary to these methods. It is not restricted to administrative data and can be applied using more commonly available data such as firm-level surveys.

Cost overreporting arises when firms acquire false invoices to claim additional tax deductions on value-added (VAT) and corporate income taxes (CIT). According to the OECD's document *Technology Tools to Tackle Tax Evasion and Tax Fraud* (2017), cost overreporting — also known as “fake invoicing”, “ghost firms”, “invoice mills”, or “missing traders”—

permeates internationally. Reports from Latin America, Eastern Europe, Asia, and Africa claim cost overreporting led to annual tax revenue losses as large as 5.6% of the GDP, for example, in Poland, 2016 (Poland’s Minister of Finance, 2018)¹

Despite its relevance, cost overreporting has been mostly overlooked by the literature. On one hand, the few studies on this evasion strategy exploit detailed administrative data (Zumaya et al. 2021, Carrillo et al. 2022). Government tax authorities restrict access to administrative data because of firms’ confidentiality concerns. On the other hand, to the best of my knowledge, no study has attempted to structurally identify cost overreporting. Unlike the case of individuals (Pissarides & Weber 1989, Paulus 2015), when it comes to corporate tax evasion, researchers have to account for an additional source of unobserved heterogeneity, productivity. Why? Because cost overreporting might be naively quantified as low productivity. Intuitively, for a given output level, high input utilization by a firm could be explained by either the amount of input the firm overreports to evade taxes or by a negative productivity shock.

To address this gap in the literature, first I formally show that ignoring tax evasion leads to downward biased productivity estimates. I then provide a new estimation strategy using production functions to jointly recover the densities of tax evasion and productivity. The intuition works as follows. In the absence of tax evasion, the first-order conditions of the firms’ cost-minimization problem let us recover the common technology, the production function. Consequently, in the presence of cost overreporting, deviations from this common technology identify tax evasion up to the measurement error. Then, from a subset of non-

¹Other reports show that cost overreporting led to revenue losses of 0.2% of Chile’s GDP in 2004 (Gonzalez and Velasquez, 2013; Jorrat, 2001; CIAT, 2008); 0.2% of Colombia’s GDP (Portafolio, 2019); and 0.03% of Mexico’s GDP in 2018 (Senado de la Republica, 2019).

overreporting firms, the strategy identifies the production function parameters and the density of the output shock. Finally, using non-parametric deconvolution techniques, I jointly recover the distributions of tax evasion and productivity.

Applying the method using firm-level data from Colombia between 1981 and 1991 —a commonly used dataset in the production function literature—, I find evidence suggesting that firms in four of the top five industries (8 of the top 20) engage in cost overreporting. These firms overreport up to 25% of their costs. My estimates suggest that the tax evading firms in the top 20 revenue industries caused the government of Colombia approximately XXXX in tax revenue losses. [exporters/importers/proprietorships/limited liability companies] are [less/more] likely to engage in cost overreporting. I also find that ignoring cost overreporting leads to consistently larger elasticities of intermediate inputs by [what factor]. The bias on the intermediate inputs spreads to the elasticities of labor and capital, whose bias direction varies by industry [what range?]. Lastly, I find significant differences in the productivity distributions. In particular, true productivity distributions are [how]. The differences between [exporters/corporations] and [importers/limited liability companies] are [what?].

1 References

Carrillo, P., Donaldson, D., Pomeranz, D. & Singhal, M. (2022), ‘Ghosting the tax authority: Fake firms and tax fraud’.

URL: <http://www.nber.org/papers/w30242>

Paulus, A. (2015), ‘Tax evasion and measurement error: An econometric analysis of survey data linked with tax records’.

URL: <http://hdl.handle.net/10419/126489>

Pissarides, C. A. & Weber, G. (1989), ‘An expenditure-based estimate of britain’s black economy’, *Journal of Public Economics* **39**, 17–32.

Technology Tools to Tackle Tax Evasion and Tax Fraud (2017).

URL: www.oecd.org

Zumaya, M., Guerrero, R., Islas, E., Pineda, O., Gershenson, C., Iñiguez, G. & Pineda, C. (2021), ‘Identifying tax evasion in mexico with tools from network science and machine learning’.

URL: <http://arxiv.org/abs/2104.13353>