The London School of Economics and Political Science

Essays in Applied Economics

Adrien Bussy

A thesis submitted to the Department of Economics of the
London School of Economics and Political Science for the degree
of Doctor of Philosophy, London, May 2020

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of approximately 39,000 words, excluding graphs, tables and appendices.

Statement of conjoint work

I confirm that Chapter 2 was jointly co-authored with Friedrich Geiecke, and I contributed half of this work.

Abstract

This thesis is composed of three chapters. Chapter 1 investigates the extent to which firms evade taxes on corporate profits by misreporting their imports and exports. I indirectly estimate evasion by studying the co-movements of tax rates and discrepancies between reports by importers and exporters of the same trade flows. I find evidence of significant evasion. To facilitate causal interpretation of the results, I develop a model of the compliance problem of the firm and test several predictions linking firms' evasion motives to accounting rules, institutional features, and the interaction of the rates of tariff and corporate income tax.

Chapter 2 focuses on innovation. We use methods from Natural Language Processing to characterize the innovative content of patents. We develop several metrics that compare inventions to existing and future innovations. The intuition guiding us is that patents dissimilar to past inventions and similar to future ones may have anticipated or started shifts in innovation topics. We find non-causal evidence that such patents have higher citations and the firms owning them grow faster and are more profitable relative to other firms. Analysis of trends suggests that innovative ideas may have gotten harder to find over time in high-innovation fields.

Chapter 3 quantifies Value Added Tax (VAT) fraud eliminated by a reform of the European Union VAT rules preventing traders from exploiting a weakness of the system that allowed them to collect VAT without remitting it to the tax authorities. The targeted fraud schemes involve cross-border transactions and result in firms misreporting their transactions, which in turn affects trade statistics. Based on 54 reform episodes, I find that fraud worth up to EUR 1.6 billion annually is removed thanks to the reform. Although large in absolute terms, this only represents a tiny fraction of total VAT revenues in the European Union.

Acknowledgements

As this venture is coming to an end, there are a number of people to whom I would like to express my gratitude. First to my supervisors for invaluable guidance. To Francesco Caselli for showing me the importance of stripping arguments and ideas down to their essence, and for having kept me motivated throughout. To Daniel Sturm for his generosity with his time and thoughts, and for his contagious enthusiasm.

To friends and colleagues who provided me with precious comments and support throughout the years of my PhD and in particular to Daniel Albuquerque, Laura Castillo-Martínez, Wouter den Haan, Thomas Drechsel, Dita Eckardt, Andreas Ek, Martina Fazio, Nicola Fontana, Ethan Ilzetzki, Xavier Jaravel, Felix Koenig, Sevim Kösem, Jay Euijung Lee, Edoardo Leonardi, Ben Moll, Rachel Ngai, Tsogsag Nyamdavaa, Łukasz Rachel, Akash Raja, Daniel Reck, Ricardo Reis, Wolfgang Ridinger, Thomas Sampson, Catherine Thomas, John Van Reenen, and Céline Zipfel. I am especially grateful to Shengxing Zhang for our many meetings and his advice, and to Dominic Rohner and Fabrizio Zilibotti for their instrumental mentorship when I was applying to the PhD programmes.

Andrea Alati, Miguel Bandeira, Patrick Coen, and Friedrich Geiecke have been true companions in this adventure. The summer leading to the job market would not have been the same without them. I thank Miguel for our countless discussions on research and for his friendship as we made our way through the PhD, and Friedrich for his inexhaustible ardour and good spirits as we worked together on chapter 2 of this thesis.

I thank my childhood friends Chloé, Léonard, Guillaume & Mélissa, Marion, Nicolas, Pierre, Pierre-Emmanuel and Vincent for always being there for me when I returned home, although we have been separated most of the time during the past six years.

Last and closest to my heart, I am grateful to my family and my godparents for their ever-renewed support. I thank my parents in particular, for sowing the seeds of intellectual curiosity throughout my childhood and for encouraging me to pursue my goals, whatever they are. I could not have reached this milestone without the unconditional and loving support of my wife, to which words cannot do justice. I like to think that this is as much my PhD as hers.

À ma femme, Qing

À mes parents

Contents

List of Figures

List of Tables

Chapter 1

Corporate Tax Evasion: Evidence From International Trade

1.1 Introduction

Revenues from taxation of corporations' profits represent a sizeable share of overall tax revenues, ranging from around 9% in OECD countries to 15% in Africa, Latin America and the Caribbean.[1] In recent years, growing concerns that significant amounts of corporate income tax are evaded, in part fuelled by large fiscal deficits following the Great Recession and instances of document leaks exposing firms' schemes aiming at dodging tax authorities, have led to considerable policy attention (Slemrod, 2016).[2] Yet the scale and the determinants of corporate tax evasion remain elusive due to its illegal and secretive nature.

I provide novel evidence of corporate income tax evasion in the form of an elasticity of *evasion* with respect to the statutory tax rate by exploiting an empirical strategy used in the literature on tariff evasion in international trade. In trade statistics, each trade flow is reported twice: once by the exporter as an export, and once by the importer as an import. The difference between these reported figures (reported exports minus reported imports) is defined as a *reporting gap*.[3] When the incentive to (mis-)report the trade flow changes for one side but remains constant for the trading partner, the reporting gap changes as a result. Fisman and Wei (2004) originally used

---

[1] Source: OECD.Stat, *Global Revenue Statistics Data Set*. The numbers are for year 2016.

[2] See, for instance, the CumEx-Files, LuxLeaks and the Paradise Papers. The Panama Papers were mostly about individuals rather than corporations.

[3] Other studies reviewed later in this section have also named it *trade gap*, or *missing trade*. The UN Comtrade unit refers to it as *bilateral asymmetries*.

this idea to study tariff evasion. Higher tariffs on imports may induce importers to lower the value of their reported imports to evade tariffs, resulting in a change in the reporting gap.

This paper extends this logic to corporate income tax (CIT) evasion. In the case of CIT, as the tax rate increases in a given country, a firm may want to under-report its exports (which are a component of sales) to lower its taxable profits. The same firm may also want to misrepresent its imports (which are a component of costs), though whether it wishes to overstate or understate is theoretically ambiguous, as I discuss below. Insofar as the trading partner's incentives remain unchanged and in the absence of confounding factors, the reporting gaps would react to changes in the CIT rate, revealing evasion. This paper estimates how reporting gaps react to tax rate changes and discusses the assumptions under which the gaps' response to these changes can be interpreted as evidence of tax evasion.

To guide the interpretation of the results, I present a model of the compliance problem of the firm. The firm acquires imported inputs and sells its products abroad, and must decide on how much of each to report to the tax authorities for the purpose of tariff and CIT taxation. Reported imports and profits constitute the tax bases for tariffs and CIT, respectively. Misreporting entails a risk, modelled in reduced form as a cost function that captures the expected cost of evasion. Deviations of the reported amounts from the true values of sales and purchases increase the cost, as do large discrepancies between reported sales and reported purchases. The latter captures the fact that a firm which reports excessively low sales and high costs may appear suspicious to the authorities. The need for consistency in the firm's accounts effectively *binds* the levels of optimal reported sales and purchases to each other through the cost function. The reporting gap in the model is constructed by considering two firms, located in different countries and subject to different levels of tax, trading one good that constitutes exports for one firm and imports for the other.

The model delivers testable predictions regarding how firms react in terms of their reported (foreign) sales and purchases to changes in the CIT and tariff rates they are subject to. First, a rise in tariffs induces the importing firm to under-report imports, which increases the reporting gap. Second, an increase in the CIT rate in the exporting country induces the exporting firm to under-report exports, leading to a decrease in the reporting gap. Third, the extent of misreporting in response to changes in CIT and tariff rates decreases when the taxman's vigilance is higher. Lastly, following an increase in the CIT rate in the importing country, whether the firm under- or over-reports will depend on two countervailing forces: on one hand, over-reporting imports would directly lower taxable profits, which is desirable for evasion purposes. On the other hand, doing so would result in the simultaneous report of low sales and high costs, thereby creating a discrepancy in the firm's accounts and increasing the likelihood of being audited. Over-reporting imports in response to an increase in the CIT rate is more likely to be optimal when (i) tariffs are high, because tariff payments are CIT-deductible and over-reporting imports lowers taxable profits through both the reported cost of purchases and the tariff liability; (ii) costs are fully CIT-deductible and therefore lower taxable profits one-to-one; (iii) the attention of tax authorities to discrepancies in terms of reported sales and purchases is low, making inconsistencies in the income statement of the firm less likely to trigger an audit. This ambiguity of the importer's behaviour offers interesting testable predictions which I fully exploit empirically.

I collect data on dyadic international trade flows between countries at a highly disaggregated product level for two separate samples: the first is an annual panel of around 63 million observations spanning 195 countries over the years 1988-2017. The second is a monthly panel dataset for EU countries containing 125 million observations over the same period.[4] The trade data is complemented with data on tariffs (over 14 million tariff changes) and corporate income tax rates (close to 500 tax changes), along with other macroeconomic and product-specific variables. The analysis of CIT evasion mostly focuses on OECD countries (268 tax changes).

The identification approach to estimating evasion primarily relies on panel regressions of the reporting gaps on the different tax rates, controlling for an extensive set of fixed effects. My identification strategy exploits *within* product-country pair variation. This choice is justified by the multitude of factors that influence the gaps, including measurement errors and cross-country differences in the way trade statistics are compiled, which increase the risk of omitted variable bias (UNSD, 2019b).

I provide novel evidence of corporate tax evasion in the form of an elasticity of

---

[4] Note that the use of aggregated trade flows as opposed to firm- or transaction-level data will pose a number of identification challenges. However, if transaction-level data *across* countries were available — which is not the case to the best of my knowledge — evasion via trade misreporting would arguably be less prevalent as tax authorities would be able to systematically cross-check each transaction based on that data.

evasion with respect to the statutory tax rate. First, the reporting gaps decrease by 0.15% to 0.3% upon a one percentage point increase in the tax rate in the exporting country, which is consistent with the prediction that exporters under-report exports. Put differently, the elasticity of the reporting gaps with respect to the net-of-tax rate faced by the exporter is 0.1. Furthermore, there is no systematic relationship between reporting gaps and the CIT rate in the importing country in the baseline estimates. A tentative calculation of evasion *levels* via these channels suggests losses in revenues for governments representing around 16% of actual corporate income tax revenues in the OECD.[5] This is broadly in line with estimates of non-compliance based on compliance gaps metrics that many countries compute nowadays (Keen and Slemrod, 2017).[6]

However, the apparent absence of co-movement between reporting gaps and the CIT rate in the importing country masks richer results. First, this relationship is positive when tariffs are low and becomes negative for higher levels of tariffs. This is consistent with the model that predicts stronger incentives to over-report imports in response to a CIT rate increase when tariffs are high, since tariff payments are CIT-deductible. Second, within-EU trade is an environment with no tariffs and where detection of inconsistencies in the firms' income statements is facilitated by the way in which Value Added Tax (VAT) returns and trade statistics are collected, which the model predicts should induce importers to under-report in response to an increase in the CIT rate. Empirically, the relationship between gaps and CIT in the importing country is positive when focusing solely on within-EU trade (consistent with importers under-reporting) and negative otherwise (consistent with importers over-reporting). Third, the association between the importer's CIT and the reporting gaps is positive for capital goods, which are typically not fully cost-deductible. This suggests under-reporting by importers, as predicted by the model.

---

[5] Given the data underlying the analysis, misreporting can only be measured for international transactions and consequently, these estimates do not reflect total evasion. Furthermore, the calculation of evasion levels presents several challenges and a discussion thereof is provided.

[6] Current estimates of taxpayers' non-compliance by tax authorities usually take the form of *compliance gaps* that compare the amount of tax due by law and that actually collected, and generally rely on two strategies: top-down and bottom-up approaches (European Commission, 2018a). The top-down approach estimates the potential tax base and revenues based on macroeconomic aggregates mostly from national accounts. The difference between potential and actual revenues gives an estimate of non-compliance but cannot differentiate between evasion and legal avoidance. The bottom-up approach relies on the outcomes from audits of individual firms, which are used to infer non-compliance and its cost at the population level. The implementation of these methods is fraught with difficulties, notably because estimating the theoretical CIT base from macroeconomic aggregates is complex (Ueda, 2018), and because audits are expensive and inferring economy-wide non-compliance is difficult when audits are not random, but assigned based on risk for instance (IRS, 2016). For a critique of the use of tax compliance gaps, see Gemmell and Hasseldine (2014).

To shed further light on these key findings, I extend the analysis in several ways. First, the negative relationship between exporters' CIT and the gaps, interpreted as exports under-reporting, decreases in magnitude when governments implement evasion-curbing policies. Second, I observe similar yet more pronounced patterns between reporting gaps and tax rates in trade in services, a notoriously fertile ground for evasion and avoidance purposes (Tørsløv et al., 2018). Third, the estimates are robust to controlling for simultaneous changes to the corporate income tax base, which was broadened significantly over the sample period in many OECD countries (Kawano and Slemrod, 2016). Last, data at the monthly and quarterly frequencies is used to explore dynamic effects that are consistent with the evasion narrative given above. In an event study framework, the reporting gaps are shown to react around the dates at which taxes change in a similar fashion as found based on yearly data. Furthermore, I uncover evidence that reported exports that appear to accrue to December are postponed to the subsequent tax year. I interpret this as evidence of inter-temporal profit-shifting by firms, and provide evidence that this pattern may be related to tax evasion.

This study also finds strong evidence of tariff evasion, echoing past studies on the topic. A one percentage point increase in the tariff rate is associated with an increase in the reporting gaps of 0.16%, reflecting a decrease in reported imports relative to reported exports. This semi-elasticity is in line with existing estimates, which range from 0.1% to 4.6% (Fisman and Wei, 2004; Mishra et al., 2008; Javorcik and Narciso, 2008; Stoyanov, 2012; Javorcik and Narciso, 2017). This pattern is more prominent the lower the income and the higher the corruption levels of the importing country. I find evasion to be higher for differentiated products for which no reference price is readily available, and to also take place through misclassification, as previously found in the literature (Javorcik and Narciso, 2008; Fisman and Wei, 2004). My results suggest that over the last 10 years, an estimated upper-bound of 44 billion US dollars worth of import duties was evaded annually across the countries in the sample — this represents 2.5% of trade subject to tariffs and about 32% of actual revenues from import duties in a subset of less-developed countries for which data on customs revenues is available.

Relation to the literature This paper contributes to three strands of the literature: first, it relates to studies that use reporting gaps to estimate tariff evasion. The idea to compare values of the same trade flow reported independently by the importing and exporting countries originates in the seminal work by Fisman and Wei (2004). Using data on trade between China and Hong Kong, the authors found evidence of tariff evasion by investigating the effect of changes in tariff rates on the reporting gaps in a cross-section of products. This method was later applied in the same spirit to study tariff evasion in different countries and times. To alleviate concerns of endogeneity of tariffs (i.e. the risk that border authorities may set tariffs based on the evasion motives of the importers), several papers have exploited quasi-experimental setups based on trade agreements that exogenously changed tariff levels (Mishra et al., 2008; Javorcik and Narciso, 2008; Stoyanov, 2012). Evidence in other studies generally support the strong and robust positive association between tariffs and evasion as measured via reporting gaps (see, amongst others, Javorcik and Narciso, 2017; Kellenberg and Levinson, 2019; Levin and Widell, 2014; Ferrantino et al., 2012; Farhad et al., 2018; Demir and Javorcik, 2020). My paper contributes to this strand of the literature in two ways thanks to the extensive coverage of the data. First, it arguably allows to better isolate the evasion component of the reporting gaps by controlling for numerous fixed effects, reducing the risk of endogeneity and omitted variable bias. Second, the extent of evasion can be related to countries' characteristics: low income, high corruption countries experience more tariff evasion.[7] To the extent of my knowledge, this is the first paper to use data on the reporting gaps at such a large scale.

Second, this paper contributes to the literature on CIT evasion and avoidance by providing an alternative and complementary strategy for estimating evasion. Besides the methods used by tax authorities described in footnote 6, academic studies have interpreted bunching of declared revenues at kinks in the tax schedule or the tax base as evidence of evasion (Best et al., 2015; Kleven and Waseem, 2013; Chetty et al., 2011; Saez, 2010), as well as used survey and audit data to study evasion, notably in relation to corruption (Alm et al., 2016; DeBacker et al., 2015; Joulfaian, 2000). Besides evasion,

---

[7] Previous literature has shown a strong yet complex link between tariff evasion and corruption. Focusing on the relation between exporters and border authorities, Dutt and Traca (2010) show that corruption impedes trade as corrupt officers extract bribes from exporters to clear the cargo through customs, yet that it may enhance trade when tariffs are high and corrupt officers allow the exporter to evade tariffs against a bribe. Sequeira (2016) explains the low trade elasticities with respect to tariffs in a quasi-experimental set up in Southern Africa by providing direct evidence on how bribes allow importers to evade tariffs, rendering them less reactive to changes in tariff rates. Both Yang (2008) and Javorcik and Narciso (2017) suggest that suppressing one evading method — namely misreporting prices — leads to the use of alternative ways to avoid paying customs duties, such as misclassifying or quantity misreporting. Not directly related to tariff evasion, Fisman and Wei (2009) show that corruption in the exporting country is correlated with the reporting gap in the markets of art and antique goods, suggesting illicit trafficking of goods undeclared in the exporting country, yet declared in the importing country where their value depends on the ability of the owner to prove their authenticity.

profit-shifting, a form of legal avoidance, has received much attention in the recent literature: multinational firms shift their profits to low-CIT or tax haven countries to reduce their tax liability, resulting in significant revenue losses for governments (Bilicka, 2019; Tørsløv et al., 2018; Dowd et al., 2017). This research also indirectly relates to the use of third-party reporting as of way of minimizing evasion. Kleven et al. (2011) show that in the context of individual taxes, evasion is more prevalent for self-reported income not subject to third-party reporting.[8] In the case of international trade, transactions with foreign entities are not directly and systematically subject to third-party reporting — as could be the case domestically if transactions are subject to Value Added Tax, for instance — and may thus be more fertile ground for evasion.

Third, this paper relates to the literature concerned with optimal corporate income tax policy, and more specifically to the estimation of the elasticity of taxable income with respect to marginal tax rates (for a review of this literature, see Saez et al., 2012).[9] This study provides an estimate of the elasticity of *evasion* with respect to the statutory tax in the context of corporate income taxation, which relates to the elasticity of taxable income (ETI). One notable difference between the two resides in the fact that ETIs encompass all behavioural responses to changes in the tax rate that lower the tax burden (reduction in effort, organisational reorganization, evasion, legal avoidance, etc.), whereas the elasticity in this paper only reflects evasion. This paper complements existing estimates of the elasticity of taxable corporate income, of which there are significantly fewer in the literature than in the case of personal income taxation (Devereux et al., 2014; Dwenger and Steiner, 2012; Gruber and Rauh, 2007).

Structure of the paper Section 1.2 presents a model of the compliance problem of the firm. Details of the data are given in section 1.3. Section 1.4 lays out the empirical strategy used to uncover tax evasion in the context of international trade. The results on tariff and corporate income tax evasion can be found in sections 1.5 and 1.6, respectively. Section 1.7 concludes. The main tables and figures can be found at the end of text, after the conclusion. Additional tables and figures, respectively referenced with prefixes 1.A and 1.B, can be found in appendices of the same name.

---

[8] Carrillo et al. (2017) and Slemrod et al. (2017) show that in the context of corporate taxation, even when taxable revenues are subject to third-party scrutiny, reporting taxpayers may adjust other margins (costs) such that taxable profits remains stable.

[9] Optimal policy is complex due to the multiplicity of the instruments available — e.g. tax base, statutory rates, enforcement level (Keen and Slemrod, 2017; Almunia and Lopez-Rodriguez, 2018) — and of the objectives to be achieved — changing tax policy not only affects tax revenues, but may also impact economic allocation and efficiency (Best et al., 2015; Dwenger and Steiner, 2014).

1.2 Model

An illustrative model is developed to show the mechanisms through which changes in tariffs and corporate income tax rates impact the reporting gaps. Some elements of the model are inspired by Ferrantino, Liu and Wang (2012). From the outset, this model should only be taken as a way to obtain testable predictions used to guide the empirical investigation in sections 1.5 and 1.6.

1.2.1 Compliance problem of the firm

Consider a firm that produces a final good using intermediate good in the production. Denote its sales F and its purchases M, both are exogenous. Given the focus on international trade, assume that all purchases are imported from abroad and are subject to tariff $t \geq 0$, and that all sales are exported.[10] Profits in this economy are taxed at rate $\tau \in [0, 1]$. Profits after tax in the absence of evasion are given by:

$$\pi^{\text{no evasion}} = (F - M) - tM - \tau \max \{0, [F - M(1 + t)\phi]\}, \tag{1.1}$$

where $\phi \in (0, 1]$ denotes the fraction of costs that are CIT-deductible. The first element of this expression is simply sales minus costs, the second is the tariff payment, and the last is the corporate income tax payment. The firm can evade both tariffs and corporate income taxes by misreporting the values of its sales and purchases. This entails a cost, as detected evasion is punished. Denoting F_r and M_r the reported amounts, the firm faces the following compliance optimization problem:

$$\max_{\{F_r, M_r\}} \quad (F - M) - tM_r - \tau \max \{0, [F_r - M_r(1 + t)\phi]\} - bC(F_r, M_r, \omega), \tag{1.2}$$

where $C(\cdot)$ is a twice continuously differentiable function that represents the expected cost of evasion, ω is a vector of parameters *not* containing the tax rates, and b is a scal-

---

[10] The assumption that all purchases are imported and all sales exported is convenient for illustrative purposes, yet obviously unrealistic. In particular, it is possible that the extent to which firms misreport in relation to taxes may differ between domestic and international transactions. Even though this does not change the predictions of the model, CIT evasion revealed through international trade may not be representative of total evasion in the economy. These considerations will be further discussed in section 1.5.3. It is possible to modify the model to encompass domestic sales and purchases. Whether the predictions change or not depends on whether firms are allowed to evade taxes by misreporting domestic transactions. If they are, a stance must be taken on whether it is easier, harder or equally hard to evade on domestic transactions. If it is equally risky to misreport domestic and international transactions, the model does not admit a unique solution. Further discussions of this issue can be found in appendix 1.C.3.

ing factor capturing how effective authorities are at detecting evasion. The expected cost depends on the reported amounts of sales and purchases. Appendix 1.C.1 contains a discussion of how the results change if the cost function is allowed to depend on the tax rates themselves, which can result in counter-intuitive results.[11] From this expression, it is clear that the tax base for tariffs is reported imports (M_r), and the tax base for CIT is reported profits. The objects of interest are how optimal reported purchases and sales that solve (1.2), denoted $M_r^\star$ and $F_r^\star$, react to changes in the tax rates t and τ. Let C_k denote the derivative of $C(\cdot)$ with respect to its kth argument, and C_{kj} the second derivative of $C(\cdot)$ with respect to its kth and jth arguments.

Proposition 1. *Under the following conditions on the Jacobian of $C(\cdot)$ at the optimal solution $(F_r^\star, M_r^\star)$:*

$$sign(C_1) = sign(-\tau), \tag{1.3}$$

$$sign(C_2) = sign(\tau(1+t)\phi - t), \tag{1.4}$$

and on the Hessian of $C(\cdot)$, denoted H_C:

$$C_{11}, C_{22} > 0, \; C_{12} = C_{21} \leq 0, \; \det\left(H_C\right) > 0, \tag{1.5}$$

the following comparative statics are signed as follows:

$$\frac{\partial M_r^\star}{\partial t} < 0, \quad \frac{\partial(F_r^\star - M_r^\star(1+t)\phi)}{\partial \tau} < 0, \tag{1.6}$$

$$sign\left(\frac{\partial M_r^\star}{\partial \tau}\right) = sign\left(C_{21} + (1+t)\phi C_{11}\right), \tag{1.7}$$

$$sign\left(\frac{\partial F_r^\star}{\partial \tau}\right) = sign\left(-(1+t)\phi C_{12} - C_{22}\right). \tag{1.8}$$

Furthermore, the magnitudes of $\frac{\partial F_r^\star}{\partial k}$ and $\frac{\partial M_r^\star}{\partial k}$ for $k = t, \tau$ decrease as b increases.

Proposition 1 — proof in appendix 1.C.1 — lays out conditions on the generic cost function such that an interior solution exists, and that allow to sign the comparative statics of interest. The conditions on the Jacobian ensure an interior solution, basically guaranteeing a solution where marginal benefits of misreporting (the right-hand sides of (1.3) and (1.4)) equal marginal costs (the left-hand sides). From these equations, it is clear that sales will be under-reported, and purchases can either be over- or under-reported depending on relative levels of CIT and tariffs: high tariffs induce the firm to under-report imports, whereas high CIT taxes induce over-reporting to reduce taxable profits, *ceteris paribus*.

---

[11] For example Yitzhaki (1974), commenting on Allingham and Sandmo (1972), shows that if tax rate increases, evasion may decreases when the penalty is based on the amounts of evaded tax.

Conditions in (1.5) are sufficient for $C(\cdot)$ to be convex. The elements on the diagonal of the Hessian are positive, which implies that the marginal costs of misreporting are increasing. The off-diagonal elements of the Hessian are weakly negative. If equal to zero, the marginal cost of misreporting sales only depends on the level of reported sales (and not on the level of reported purchases) and conversely for reported purchases. If negative, the reported level in one dimension affects the marginal cost of misreporting in the other dimension. More specifically, $C_{12} < 0$ *binds* the reported levels of sales and purchases to each other through the cost function in the following intuitive way. A firm reporting excessively high purchases and low sales looks suspicious to the tax authorities.[12] Therefore, reporting low purchases leaves more room to report low sales without looking suspicious. Similarly, reporting high sales leaves more leeway to report high purchases.[13] Last, the determinant of the Hessian must be positive, such that the interdependence between reported sales and purchases does not overpower the basic shape of the marginal costs.

Under the conditions in proposition 1, optimal reported purchases decrease as the tariff rate increases, and reported taxable profits decrease as the CIT rate increases — see (1.6). These two effects reflect the desire of the firm to lower the reported tax base when the rate increases, for both tariffs and CIT. Remains the question: how does the firm manipulate $F_r^\star$ and $M_r^\star$ to lower taxable profits? In principle, the firm could either: (i) lower reported sales and inflate reported costs; (ii) simultaneously lower reported sales and reported costs (but by less); or (iii) simultaneously inflate reported sales and reported costs (but by more). In all cases, reported profits decrease.

Assumption 1. $-(1+t)\phi C_{12} - C_{22} < 0.$

Assumption 1 rules out case (iii) above. Intuitively, it is easier to imagine a firm under-reporting its sales and purchases, e.g. by hiding part of its business, rather

---

[12] In the context of taxation of corporate profits of Nicaraguan firms in 2011-2012, Carrillo et al. (2017) show that in response to notifications of discrepancies in self-reported income by tax authorities to firms, firms that react to the notification do so by declaring more revenues, yet simultaneously increase reported costs. Slemrod et al. (2017) find a very similar pattern in the case of small-business tax compliance, when the US Internal Revenue Service gained access to credit-card information on electronic sales as a source of third-party reporting. Naritomi (2019) finds similar patterns in the case of VAT in Brazil.

[13] To give a detailed explanation, consider how the marginal cost of under-reporting sales (C_1) changes as the level of reported purchases increases. First, notice that the marginal benefit or misreporting sales is $-\tau$, so the firm under-reports sales to save on CIT. The cost increases as it does so, since $C_1 < 0$. The more it under-reports, the higher the marginal cost as $C_{11} > 0$. Because $C_{12} < 0$, the marginal cost of under-reporting sales increases when reported purchases are higher as reporting low sales looks suspicious in that case. A similar intuition can be given when the firm over-reports purchases ($C_2 > 0$). Since $C_{22} > 0$, the marginal cost of over-reporting purchases increases the more the firm over-reports. Because $C_{21} < 0$, the higher reported sales are, the more room it leaves for the firm to over-report purchases, since high reported sales make reporting large amounts of purchases less suspicious.

than creating imaginary sales and purchases for the purpose of evasion. Under this assumption, optimal reported sales decrease in response to an increase in the CIT rate (τ). The last comparative statics of interest is that of reported costs with respect to τ. Lowering reported taxable profits can be achieved by lowering reported sales and inflating costs. As explained above, this urge is counteracted by the fact that high reported costs and low reported sales will appear suspicious to the taxman. This could lead the firm to lower reported sales and purchases simultaneously (whilst ensuring that reported profits decrease overall). In that case, one can think of the firm reporting in such a fashion as to appear smaller than it actually is. This happens when

$$(1+t)\phi C_{11} < -C_{21}. \tag{1.9}$$

This condition is more likely to hold when tariffs are low because import duties are deductible from the CIT bill and high tariffs therefore encourage over-reporting imports as both the cost of materials and tariff payments lower taxable profits. Condition (1.9) is also more likely to hold when purchases cannot be fully deducted from taxable profits (i.e. low ϕ), which renders over-reporting less effective at lowering reported taxable profits, and for high values of $|C_{21}|$, which bind the levels of $F_r^\star$ and $M_r^\star$ more tightly to each other. In that case, an increase in τ induces the firm to under-report purchases ($M_r^\star$) in tandem with the decrease in reported sales ($F_r^\star$).[14] The analysis above is repeated in appendix 1.C.4 with an explicit illustrative and more intuitive cost function that admits closed-form solutions.

1.2.2 Implications for the reporting gaps

The previous section detailed the problem of a generic firm in a given country. To link the evasion behaviour of the firm above to the reporting gaps, consider two such firms: an importer (I) and an exporter (E) located in different countries. All variables above are indexed by the superscript $k \in \{I, E\}$ to indicate to which side they refer. Each firm solves the problem above, facing corporate tax rates and tariffs (τ^k, t^k). Furthermore, consider the case where the good exported is used as intermediate input in the production of the importer, i.e. $F^E = M^I$. The reporting gap for the flow between

---

[14] The last term in (1.9) is C_{11}, which cannot be too large for the condition to hold. This is because for reported purchases to react negatively to τ, reported sales must decrease *enough*, which will not be the case if C_{11} is very large (because the marginal cost of under-reporting sales would increase too rapidly as the firm under-reports more).

these two firms is thus:

$$\text{Reporting gap} = \log(F_r^{E\star}) - \log(M_r^{I\star}). \tag{1.10}$$

The difference in logs is not central to the predictions — i.e. a simple difference would leave the predictions unchanged — but corresponds to the measure of the gaps that will be used in the empirical section. From proposition 1, the main testable predictions from the model immediately follow.

Testable predictions. *The reporting gap responds to changes in the tax rates in the following way:*

1. *Upon an increase in the tariff rate in the importing country (t^I), reported imports ($M_r^{I\star}$) decrease, which leads to an increase in the reporting gap.*
2. *Upon an increase in the corporate income tax rate in the exporting country (τ^E), reported exports ($F_r^{E\star}$) decrease, which leads to a decrease in the reporting gap.*
3. *When condition (1.9) holds, an increase in the corporate income tax rate in the importing country (τ^I) induces the importer to lower reported imports ($M_r^{I\star}$), which leads to an increase in the reporting gap; and conversely when the reverse of (1.9) holds. Condition (1.9) is more likely to hold when:*
    - *(a) Tariffs are low (t is low);*
    - *(b) Costs are not fully CIT-deductible (ϕ is low);*
    - *(c) Tax authorities closely monitor inconsistencies in income statements ($|C_{12}|$ is high).*
4. *The reactions of the reporting gap to changes in CIT and tariffs are dampened when the authorities are better at detecting evasion (b is large).*

The main takeaway from this section is that the reaction of the reporting gap with respect to the CIT rate faced by the importer is not obvious. It can increase or decrease depending on parameter values. This will be particularly relevant for (i) trade within the European Union, where tariffs are zero and I will argue that $|C_{12}|$ is particularly high, making it more likely that reported imports decrease in response to an increase in the corporate income tax rate; and (ii) capital goods for which ϕ is typically smaller than for other goods, since capital purchases must be accounted as assets and depreciated gradually over time.

1.3 Data

This section provides details on the data, as well as descriptive statistics of the main variables. There are two samples. The first one is a panel of 195 countries over 1988-2017 at a yearly frequency. The second one is panel limited to trade between countries of the European Union over 1988-2017, at a monthly frequency. These will be referred to as *world sample* and *within-EU sample*, respectively.

1.3.1 Sample definition

World sample The world sample consists of a panel of 195 countries over the years 1988 to 2017. The unit of observation is a product at the Harmonized System (HS) 6-digit level for an importer-exporter pair — an (i, e, p) triplet — for which the reporting gap is observed over time. The sample is strongly unbalanced, as more countries and products enter the sample over time. Even conditional on being observed once, an (i, e, p) triplet may disappear and reappear over time. I re-estimate all regressions on a more balanced sample for robustness.

Within-EU sample The within-EU sample consists of all EU member states over the period 1988-2017. The sample expands as more countries joined the EU, from 12 in 1988 to 28 in 2017. Similarly to the world sample, (i, e, p) triplets may be missing for some time periods before re-entering the sample. Unlike in the world sample, observations are observed monthly.

The within-EU sample is used for an event study, which relies on accurate measurement of the time at which a new corporate income tax rate becomes binding for firms. The fiscal year coincides with the calendar year in every EU country except the UK, where the fiscal years starts in April. However, firms in most countries are not legally forced to choose the fiscal year as their financial year for accounting proposes (e.g. firms may choose September as their year-end month). The relevant tax rate that applies to a firm whose accounting year does not coincides with the fiscal year varies across countries.[15] Given that only aggregate data is available, it effectively renders

---

[15] For instance in Germany, the tax rate to be applied is that which is valid on the date at which the

impossible dating the moment at which new tax rates become binding when accounting periods vary across firms within a country. In an effort to address this issue, firm level data from Bureau Van Dijk's (BVD) Amadeus is used to compute the fraction of firms whose accounting period coincides with the fiscal year in their country of incorporation. The results can be found in figure 1.B.5: even though in some countries a significant number of firms do not choose December as their year-end month, these firms are not economically important when measured in terms of total assets (or revenues).[16]

1.3.2 Trade data and reporting gaps

Trade statistics come from the United Nation Comtrade database and Eurostat.

1.3.2.1 World trade data

UN Comtrade gathers international trade statistics from national authorities and makes them centrally available (UNSD, 2017). For each country pair and product, Comtrade provides the trade flow type (imports or exports), the trade value and the quantity traded. Trade flows are available at the 6-digit level of the Harmonized System (HS), that comprises about 5,300 products. Flows are reported gross. Comtrade also provides information on re-imports — "goods imported in the same state as previously exported" — and re-exports — "exports of foreign goods in the same state as previously imported" (*UNSD*, 2019a). To be counted as re-imports or re-exports, the goods cannot have been processed during their time abroad.

Collection of trade statistics by Comtrade Trade values are converted into current US dollars by Comtrade when the original data is in another currency. Trade values are generally checked for major inconsistencies (e.g. negative values, totals very far from official aggregate statistics), and classified into normalized commodity codes — the Harmonized System — when national data items are otherwise classified. Imports are generally recorded with more accuracy than exports, as they are subject to tariffs and thus subject to more scrutiny by customs.

---

accounting year of the firm ends. In the UK, if a tax change occurs during the accounting year of a firm, that firm must apply both the old and the new rates in proportion of the overlaps between the accounting period and the two fiscal years that it spans (source here, last accessed 18/05/2020).

[16] The only exception is the UK, where only 20% of (asset-weighted) firms have their year-end in March. All results in the event-study are also estimated dropping the UK from the sample for robustness.

Data on quantities may be of less good quality. There are more missing values than in the case of trade values — which are usually the relevant metrics for the calculation of duties. Furthermore, quantities are reported in different units (e.g. tons, units, meters). When quantities are expressed in non-standard units (as defined by the World Customs Organization), Comtrade converts them into standard units. When quantities are not available, Comtrade estimates them when possible "using the reporter's reported weighted unit value within the same 6-digit commodity flow or median standard unit value of all reporters for the same 6-digit commodity flow the previous year, after elimination of outliers" (see p.11 of UNSD, 2017). Those two facts seem to suggest that quantities may be more prone to suffer from measurement errors than trade values and will therefore receive less weight when interpreting regression results.

1.3.2.2 EU trade data

EU data on international trade comes from Eurostat and is available since 1988 at a monthly frequency. For every importer-exporter pair, Eurostat provides the value and the quantity traded at the product level. The trade flows are highly disaggregated, down to the 8-digit level of the Combined Nomenclature (CN) — the coding system used in the EU to classify traded goods.[17] There are circa 10,000 products at the 8-digit level. All flows are reported gross. The analysis will focus on flows aggregated to the HS 6-digit level to reduce noise in the reporting gaps, and because of computing power limitations. For the same reasons, data will be aggregated to the quarterly frequency in some cases.

Collection of trade statistics in the EU Since the creation of the Single Market in 1993, goods traded between EU members do not have to clear customs, the traditional source of trade statistics. Trade data is directly provided by traders, who report detailed information on their transactions to their national authorities via a system called Intrastat. Member States then forward the data to Eurostat, which compiles trade statistics for the whole bloc. Intrastat is closely tied to the VAT system. The formats of Intrastat declarations and VAT returns are similar, as both require traders to report their total intra-EU imports and exports. This allows authorities to check for the accuracy of Intrastat declarations. Although Intrastat declarations are not payoff-

---

[17] The six first digits of the CN coincide with the Harmonized System (HS), the nomenclature developed by the World Customs Organization and used in other trade databases, such as UN Comtrade.

relevant for firms — in the sense that no tax is based on them — firms have a strong incentive to be accurate as discrepancies may be investigated and fined.[18]

In order to reduce the administrative burden on small and medium-sized businesses, firms whose yearly imports or exports are below certain thresholds are exempted from filling Intrastat declarations. These thresholds are set independently by each country and can change every year, with effect on January 1st. They are currently set such that at least 97% of total dispatches and 93% of total arrivals are covered by Intrastat.[19] In the past, coverage requirements were higher, implying lower thresholds. The coverage requirements were significantly lowered in 2004.[20] Figure 1.B.3 displays the average and medium thresholds over time across EU members. Figures 1.B.1 and 1.B.2 plot the time series of the thresholds by country for imports and exports, respectively.[21] A remarkable number of firms are exempted from Intrastat declarations: in 2017, an average of 12.6% and 20.2% of VAT-registered traders were providers of statistical information for imports and exports, respectively (Eurostat, 2017).[22] This suggests that international trade is highly skewed towards large firms.[23] The remaining 3% of exports and 7% of imports are estimated by the national authorities based on the VAT returns of the firms falling below the reporting thresholds. The within-EU sample excludes these estimates and only contains trade reported by firms above the Intrastat thresholds. The implications for the reporting gaps and for the estimation of evasion will be discussed in section 1.3.2.4. I control for the levels of the thresholds in all regressions based on the within-EU sample.

1.3.2.3 Constructing the reporting gaps

The reporting gaps are constructed using mirror statistics, i.e. information on the same gross flow that is reported twice: once by the importer (as an import) and once by the exporter (as an export).[24] The gaps are computed in terms of value and quantities.

---

[18] See here for the UK, sections 7 and 8 (last accessed 18/05/2020).

[19] EC regulation No 222/2009, art. 6(a), and Eurostat (2016, section 6.2.4).

[20] EC regulation No 638/2004, art. 10(3). This law replaced Council Regulation (EEC) No 3330/91 and specified a coverage of 97% for both imports and exports.

[21] The data on thresholds was manually collected from a variety of sources, and can be found online at the following Github repository.

[22] See table 1.A.1 for a breakdown by country.

[23] If tax evasion were more prevalent among smaller firms — which seems likely as larger firms are audited regularly and may prefer legal tax avoidance over outright evasion — the fact that large firms are responsible for most of the trade volume implies that estimates of evasion based on aggregated trade flows is probably a lower-bound for evasion among smaller firms.

[24] Net flows (e.g. net imports are imports minus exports) for a given product and country pair should be the mirror image of each other, since net imports of a country corresponds to net exports of the partner

Formally:

$$X\ \text{gap}_{iept}^{log} = \log(\text{exports } X_{ipt}^{e}) - \log(\text{imports } X_{ept}^{i}), \tag{1.11}$$

where $X \in \{\text{value}, \text{quantity}\}$, i denotes the importer, e denotes the exporter, p the product and t time. Superscripts denote who reports the trade flow and subscripts designate the trade partner. This convention is used throughout the paper. As an illustrative example, the value of exports of *uncooked pasta, not stuffed or otherwise prepared* (HS 190211) from Germany to France as reported by Germany, is compared to the value of imports of that same good as reported by France. The difference in logs is the measure commonly used in the literature.[25] If the gap is small, this measure approximates the difference between exports and imports as a fraction of the value of imports. For robustness, I also consider an alternative measure of the gaps that does not omit zeros:

$$X\ \text{gap}_{iept}^{robust} = \frac{(\text{exports } X_{ipt}^{e} - \text{imports } X_{ept}^{i})}{\frac{1}{2}(\text{exports } X_{ipt}^{e} + \text{imports } X_{ept}^{i})} \in [-2, 2]. \tag{1.12}$$

Flows where one side is zero will appear as a 2 or -2. Both in Comtrade and Eurostat, missing values indicate that the data is unavailable and not necessarily that trade did not take place. For this reason, observations where one side of the reporting gap is unavailable are removed from the sample. In order to minimize the risk of data errors, the bottom and top percentiles by importer in terms of the reporting gaps are removed from the sample.[26]

1.3.2.4 Known sources of bilateral asymmetries

Bilateral asymmetries — i.e. reporting gaps — are a known feature of trade statistics. There are many potential causes to it, which are often discussed among users of official statistics (see, e.g. Markhonko, 2014; UN, 2019). Most of these factors are not related to evasion motives (Braml and Felbermayr, 2019).

First, the statistical valuation of imports and exports differ: imports are reported CIF (including Cost of Insurance and Freight) and exports FOB (free on board, net of carriage costs), i.e. the reporting gap includes the cost of freight and insurance.

---

country.

[25] See references in the introduction.

[26] Similar to Javorcik and Narciso (2017), for example. As a robustness check, the results will also be computed without removing outliers.

Second, there may be differences in partner country attributions: in general, recommendations by the UN Statistical Commission regarding partner attribution stipulates that for imports, the country of origin must be recorded, whereas for exports, the last known destination country should be recorded. The application of these principles can be challenging, e.g. in the case of exports when the immediate destination country is only a step in the transit of the goods towards their final destination.[27] Particular attention must therefore be paid to re-imports and re-exports. The partner country is the country of origin when accounting for imports. Therefore in the case of re-imports, it appears as if a country imports from the country of origin, and not from the consignment country — i.e. the country from which the goods physically come from in that transaction. However, this trade flow will be counted as an export in the exporting country, creating a discrepancy (see Liu, 2013, footnote 1). In this paper, the gross flows are used — i.e. including re-imports and re-exports.[28] Third, differences in the trade system used by the reporting countries may result in asymmetries. Some countries apply the general trade system, in which the statistical territory coincides with the economic territory, i.e. it includes warehouses, free zones and free circulation areas. Other countries use the special trade system, wherein customs warehouses and other free zones are excluded from the statistical territory. Lastly, misclassification across product categories (more relevant for very fine trade statistics), differences in the exchange rate used by each side and time lags between the moment imports and exports are recorded by the respective countries may all result in asymmetries.

Sources specific to the EU For within-EU trade, some of the sources above are less relevant as trade statistics collection is largely determined by rules applicable to all EU states. The trade system for intra-EU trade is close to the general trade system and is common to all Member States (Eurostat, 2016, p.15). Similarly, the rules that must be followed to determine the partner country are identical for all EU countries. There are however other sources of discrepancies specific to intra-EU trade. First, since the data is reported monthly, it is more likely that a given transaction may be recorded in different time periods by each side. Second, as explained in section 1.3.2.2, the

---

[27] A possible solution to this problem being discussed by the UN is the use of country of consignment — i.e. the immediate country from which imports arrive and to which exports are dispatched — as opposed to the use of origin/final destinations (UN, 2011).

[28] This choice is made because data on re-imports and re-exports is missing for most observations (more than 94% of the observations). Using either the gross flows, or the flows net of re-imports and re-exports yields almost identical results. The correlation between the reporting gaps constructed using each definition of trade flows is around .98.

TABLE 1.1: DESCRIPTIVE STATISTICS OF THE MAIN VARIABLES

	World sample							
	N	Mean	Std	Min	25%	50%	75%	Max
Value gap, in log	91,245,036	0.03	1.82	-9.65	-0.81	-0.01	0.87	8.02
Quantity gap, in log	79,023,322	0.01	2.18	-11.47	-1.00	0.00	1.03	11.92
Value gap, robust	91,245,041	0.01	1.10	-2.00	-0.77	-0.01	0.82	2.00
Quantity gap, robust	80,350,339	0.01	1.20	-2.00	-0.94	0.00	0.97	2.00
Tariff rate importer	64,312,349	4.35	10.58	0.00	0.00	0.00	5.00	3,000.00

	Within-EU sample							
	N	Mean	Std	Min	25%	50%	75%	Max
Value gap, in log	145,249,500	0.05	1.55	-5.85	-0.53	0.02	0.69	6.68
Quantity gap, in log	122,463,910	0.04	1.62	-6.26	-0.62	0.00	0.73	6.40
Value gap, robust	145,517,228	0.05	0.98	-1.99	-0.52	0.02	0.67	2.00
Quantity gap, robust	139,944,998	0.03	1.18	-2.00	-0.76	0.00	0.86	2.00

Note: descriptive statistics for the main measures of the reporting gaps, and of tariffs. Data from the world and within-EU samples.

existence of Intrastat thresholds will results in asymmetries, e.g. if one firm is above the relevant reporting threshold and the partner firm is not.

1.3.2.5 Descriptive Statistics

Reporting gaps in the world sample The world sample contains circa 93 million observations for which the reporting gaps are available, 60 million of which are matched to tariffs. The number of observations per year increases over time (see figure 1.B.6). Trade flows for which the reporting gaps can be constructed represent circa 77% of yearly world trade on average over the period 1993-2017.[29] Table 1.1 reports descriptive statistics of the main variables, and figure 1.B.7 displays the histograms of the reporting gaps. The mean gap (in log) varies between 0.01 and 0.03 depending on the measure (value or quantity), which approximately represents between 1 to 3% of the value of the underlying trade flow. Striking is the existence of very large gaps, and the symmetry of the gaps in general: there are roughly as many positive gaps (reported exports exceed reported imports) as negative gaps. This confirms that evasion is most likely not the main driver of these asymmetries. In aggregate, the total reporting gap, defined as the difference between the sum of reported imports and exports (across products and country pairs) and calculated as in expression (1.12), fluctuates over time and represents between -2% and 2% of aggregate trade — see figure 1.B.9.[30]

---

[29] In years 1988-1993, the coverage is much poorer, 18% on average.

[30] Aggregate trade only includes flows for which information by both importers and exporters is available. The reporting gap is larger and more volatile in the early years of the sample.

Reporting gaps in the within-EU sample Basic descriptive statistics of the reporting gaps in the within-EU sample can be found in table 1.1, and their histograms are displayed in figure 1.B.8. The gaps are similarly distributed as in the world sample, but are less noisy (smaller standard deviations) and have less mass in the tails (less extreme values), which may reflect the higher degree of harmonization of trade statistics collection and processing across EU states. The reporting gaps are lower in December and higher in January on average — see figure 1.B.10. This would be consistent with firms manipulating their accounts in an effort to lower current profits in December — the year-end for most firms — by pushing some of their exports (sales) made in the current year to the next year and possibly also bringing forward some of their imports (costs) that accrue to the coming year to the current year. This hypothesis will be further discussed in section 1.5.5. This cyclical property is also visible in the time series of the mean and median reporting gaps, as shown in figure 1.B.11.

1.3.3 Tariffs

Data on *ad valorem* tariff rates — i.e. expressed in percent of the dutiable trade value — come from the UNCTAD TRAINS database, accessed via the World integrated Trade Solutions (WITS) website. For a given importer-product-year observation, there may be one general non-discriminatory rate — called the Most Favoured Nation tariff (MFN) for WTO countries — and one or several preferential rates stemming from trade agreements. I follow the WITS instructions and take the lowest rate as the effectively applied tariff. There are no tariffs on trade within the EU. The analysis is restricted to tariffs expressed *ad valorem* and excludes non-tariff barriers to trade.[31]

Descriptive statistics In the world sample, the average tariff is 4.35% — 2.2% if weighted by trade value. More than half of the observations have zero tariff due to the many trade agreements and the EU. Some tariffs are in excess of 100%, reflecting trade punishments or other political agendas. The most extreme example is the tariff imposed on alcoholic beverages by Egypt, which reaches 3000%. Figure 1.B.15 shows the histograms of tariff rates, of the number of tariff changes per product-country pair triplet, and of the values of tariff changes. As already mentioned, tariffs are zero for most observations. For many (i, e, p) triplets, there was no change in tariffs over

---

[31] Studying non-tariff measures (such as quotas, labelling restrictions, etc.) requires computing an *ad valorem* equivalent, which is usually estimated from the trade data directly (see, e.g. Kee and Nicita, 2016).

the sample period. The variation in tariffs within country pair-product groups comes from those groups for which at lease one tariff change took place. Most changes were of sizes between -10% and 10%, and tariffs decreases are relatively more frequent than increases. There are about 14 million tariff changes in the sample.

1.3.4 corporate income tax rates

Data on corporate income tax rates comes from the OECD tax database, complemented by data from KPMG since 2003 for non-OECD countries. The variable of interest is the central government corporate income tax rate — by opposition to sub-central/regional taxes that exist in many countries. Sub-central taxes are not taken into account in the main analysis, because they only affect a subset of firms in a country and it is difficult to assign trade flows to regions within a country. When several tax brackets exist, the top marginal rate is used.[32]

The OECD provides data on both the central CIT and the combined CIT — which takes into account representative sub-central/regional taxes in that country.[33] KMPG only provides a measure of the combined CIT. The combined CIT will be used to extend the sample beyond OECD countries as a way of checking the robustness of the results. When measures of the combined CIT are available from both KPMG and the OECD, the OECD data is preferred although both measures are very highly correlated.

Corporate taxes in the EU The rules governing corporate taxation are set at the national level and can be different for each EU member. Two main elements underlie corporate taxes: the tax rate and the tax base. At present, neither of them is harmonized at the EU level, although there has been recent attempts at elaborating EU-wide rules — none of which have been implemented at the time of writing.[34]

In order to carry out the event study analysis, the exact dates at which tax changes take place are necessary. Each tax change was manually dated in order to ensure accuracy.[35] The event date is defined as the latest of the date at which the new tax rate

---

[32] As of 2018, 19 out of the 28 EU states have a flat corporate income tax rate. See table 1.A.2. Furthermore, since international trade is skewed towards large firms, the top tax bracket is likely to be the one binding for many firms from which trade originates.

[33] The exact definition of the sub-central taxes varies across countries, and can be a combination of several separate taxes (OECD, 2019).

[34] In October 2016, the European Commission proposed the Common Consolidated Corporate Tax Base (CCCTB), a single set of rules to calculate the taxable profits of firms operating across borders within the EU. It aims at implementing it from January 2020 and thus does not impact the current study. Source here (last accessed 18/05/2020).

[35] Sources for the dates of tax change are mainly national laws, as well as the EY Worldwide Corporate

TABLE 1.2: DESCRIPTIVE STATISTICS OF CORPORATE INCOME TAX RATES

	N	Mean	Std	Min	50%	Max	N min.	N max.
CIT, OECD	967	28.17	8.90	8.50	28.00	56.00	23	36
Combined CIT, World	2,290	26.02	10.99	0.00	28.00	60.10	21	170
CIT, EU	596	27.73	9.65	9.00	28.00	56.00	12	28

Note: Descriptive statistics of the main corporate income tax measures. The first line refers to the central CIT for OECD countries, the second line to the combined CIT the the world sample and the third line to the within-EU sample. The columns N minimum and N maximum refer to the minimum and maximum number of countries for which data is available across years.

TABLE 1.3: DESCRIPTIVE STATISTICS OF CORPORATE INCOME TAX RATES CHANGES

	N	Mean	Std	Min	50%	Max	$N > 0$	$N < 0$	$\text{Mean} > 0$	$\text{Mean} < 0$
Δ CIT, OECD	268	-2.21	4.05	-25.00	-2.00	10.00	53	215	2.51	-3.37
Δ Comb. CIT, World	497	-1.99	5.20	-40.00	-1.33	25.00	122	375	2.89	-3.57
Δ CIT, EU	152	-1.93	3.48	-16.20	-2.00	10.00	27	125	2.86	-2.96

Note: Descriptive statistics of the tax changes for the main corporate income tax measures. The first line refers to the central CIT for OECD countries, the second line to the combined CIT the the world sample and the third line to the within-EU sample. The last four columns indicate the number of negative and positive tax changes, and their respective means.

became binding and the date at which firms became aware of it. For example, if the new tax rate is enacted by law on March 23rd but becomes binding retroactively from January 1st of the same year, the event date is March 23rd. An illustrative diagram can be found in figure 1.B.4.

1.3.4.1 Descriptive statistics

Summary statistics of the corporate income tax rates and changes thereof can be found in tables 1.2 and 1.3, respectively. Histograms of the CIT values, changes and number of changes per country can be found in figure 1.B.12 for the OECD, world and within-EU samples. The coverage of the data generally increases over time. For the OECD and within-EU samples, this is due to countries joining these groups over time. In the case of the combined CIT, this is mainly because the KPMG data is only available since 2003. There were 268 tax changes in the OECD sample, 497 in the world sample (using the combined CIT) and 152 in the within-EU sample. The majority of the tax changes are negative, reflecting the general downward trend in corporate income tax rates across the world and in developed countries in particular. Note that for the analysis of CIT evasion, variation in the CIT rate is the relevant dimension.[36]

Tax Guide (published yearly in January). The OECD tax database provides the rate applicable as of January 1st of a given year but does not specify the dates of tax changes. This data is available upon request.

[36] The 60 million observations of trade flows may be misleading as these are not independently and identically distributed. CIT rates vary at the country level over time and thus at a higher level than that of observations.

1.3.5 Other data

GDP figures are from the World Bank. A country-level corruption index is provided by the International Country Risk Guide (ICRG). Rauch (1999) provides a measure of the extent of differentiation of products. Rauch's differentiation index classifies products into three categories: goods traded on an organized exchange (denoted w in his nomenclature), reference-priced goods (r) and differentiated products (n). Data on changes to the corporate tax base in OECD countries was kindly provided by Kawano and Slemrod (2016).

1.4 Empirical framework

This section details the identification assumptions and the regression equations that I estimate in the data. The general strategy consists in regressing reporting gaps on the different tax rates, and interpreting movements in the gaps associated with tax changes as evidence of evasion. The world sample — which contains yearly data — is mainly used to estimate panel regressions with extensive fixed effects. I investigate both tariff and corporate tax evasion. The within-EU sample — available at a monthly frequency — is used in an event-study setting to examine how gaps change around a precisely dated corporate income tax rate change, and to investigate profits shifting across years.

1.4.1 Decomposing the reporting gaps

Evasion is not the only, nor the main component of the reporting gaps.[37] Specifically, measurement errors due to differences in data treatment across countries and in statistical valuation of imports and exports (CIF versus FOB, respectively) may account for a significant fraction of the gaps. For illustration purposes, suppose that all these factors are proportional to the true trade value:

$$\begin{aligned}\text{Value gap}_{iept} = {} & \log(\text{true trade value}^{e}_{ipt} \times \text{exporter's evasion}^{e}_{ipt} \times \text{measurement error}^{e}_{ipt}) \\ & - \log(\text{true trade value}^{i}_{ept} \times \text{importer's evasion}^{i}_{ept} \times \text{measurement error}^{i}_{ept} \\ & \times (1 + \text{freight cost}^{i}_{ept})).\end{aligned}$$

---

[37] See section 1.3.2.4 above for more details.

Denoting the log of measurement errors as ξ, freight cost as F and realizing that the true trade value is identical irrespective of who reports it, the gap can be rewritten as

$$\text{Value gap}_{iept} = \underbrace{\log\left(\frac{\text{exporter evasion}^{e}_{ipt}}{\text{importer evasion}^{i}_{ept}}\right)}_{\text{Gap due to evasion}} + \underbrace{\xi^{e}_{ipt} - \xi^{i}_{ept} - \log(1 + F^{i}_{ept})}_{\equiv \chi_{eipt}}, \tag{1.13}$$

where χ_{eipt} regroups the components of the gap that do not pertain to evasion. The challenges of using the reporting gaps to identify evasion become apparent in this expression. Not controlling for the determinants of χ_{eipt} in the regression would result in an omitted variable bias, unless χ_{eipt} is uncorrelated with the CIT and tariff rates. Identification thus relies on fixed effects and controls to account for the components of χ_{eipt} that may be correlated with the tax rates.

1.4.2 Regression equations

The main empirical strategy consists in regressing the reporting gaps on statutory tariff and corporate income tax rates.

1.4.2.1 Panel regressions

An individual observation in the panel is defined as an importer-exporter-product triplet, which is observed over time.

Tariff evasion In order to detect tariff evasion, the following regression is estimated:

$$Y_{iept} = \alpha\, \text{tariff}^{i}_{ept} + \text{FE}_{iep} + \text{FE}_{iet} + \epsilon_{iept}, \tag{1.14}$$

where the dependent variable — Y_{iept} — is a measure of the reporting gaps as introduced in section 1.3.2.3. The parameter of interest is α, which is expected to be positive as a higher tariff is predicted to increase the gaps through lower reported imports by the importer (cf. testable prediction 1 in the model section). Note the very stringent set of fixed effects (denoted as FE above). Importer-exporter-product fixed effects are included, ensuring that only variation within groups is used. Furthermore, importer-exporter-time fixed effects absorb country pair-specific time variation.[38] This is the

---

[38] Each individual dimension composing a fixed effect is subsumed by the combined fixed effect. E.g. the *iet* fixed effect is equivalent to including the following set of fixed effects: i, e, t, et, it, ei, eit.

strictest specification. Other specifications will be tested with less stringent fixed effects and more country- or product-specific controls.

Corporate tax evasion To investigate corporate tax evasion, a similar regression is estimated:

$$Y_{iept} = \alpha\, \text{tariff}^{i}_{ept} + \beta^{imp}\text{tax}_{it} + \beta^{exp}\text{tax}_{et} + \text{FE}_{t} + \text{FE}_{iep} + \delta_{i}t + \delta_{e}t + \gamma\, \textbf{controls}_{eit} + \epsilon_{iept}. \tag{1.15}$$

The coefficients of interest are β^{imp} and β^{exp}, which capture how the reporting gaps vary in relation to the corporate income tax rates. The model predicts that an increase in the corporate income tax rate of the exporter results in a decrease in the reporting gaps through lower reported exports — cf. prediction 2 — so β^{exp} is predicted to be negative. The sign of β^{imp} depends on whether condition (1.9) of the model holds, which will be indirectly tested in section 1.5.2. Note that in addition to fixed effects, country-specific time trends are added as additional controls. Results with and without these trends will be reported.

Identification of CIT evasion is more challenging since the country pair-time fixed effects are dropped from the regression to allow for the inclusion of the corporate income tax rates, which vary at the country level over time. Controls include country memberships to the WTO and the EU, measures of corruption and real GDP for each country (total and per capita).[39] Alternatively, separate regressions are estimated for the importer and the exporter sides separately, wherein the tax rate of one side only is included as regressor (together with controls), and a country-time fixed effect can be added to control for confounding factors in the partner country. For instance, to study corporate tax evasion by the exporter, the following regression is estimated:

$$Y_{iept} = \alpha\, \text{tariff}^{i}_{ept} + \beta^{exp}\text{tax}_{et} + \text{FE}_{iep} + \text{FE}_{it} + \delta_{e}t + \gamma\, \textbf{controls}_{eit} + \epsilon_{iept}, \tag{1.16}$$

where importer-time fixed effects are included. An analogous equation can be specified to estimate tax evasion by importers:

$$Y_{iept} = \alpha\, \text{tariff}^{i}_{ept} + \beta^{imp}\text{tax}_{it} + \text{FE}_{iep} + \text{FE}_{et} + \delta_{i}t + \gamma\, \textbf{controls}_{eit} + \epsilon_{iept}. \tag{1.17}$$

---

[39] I also check that the results hold controlling for GDP growth and its lag.

All the regressions above are also estimated using the log of the net-of-tax rate, $\log(1 - \text{tax})$, a measure commonly used in the literature (see, e.g. Fuest et al., 2018).

1.4.2.2 Event Study in the within-EU sample

Since the removal of the country-time fixed effects in equation (1.15) may pose an identification problem, an event study framework is proposed. Intuitively, one looks at how the reporting gaps react to a change in the corporate income tax rates in the time periods — months or quarters in this case — neighbouring the tax change. Focusing on the within-EU sample has two advantages: first, data is available at a monthly frequency, which allows to study the effects of a tax change on impact; second, within-EU trade is not subject to tariffs, which allows the study of changes in the corporate income tax rate exclusively — tariffs cannot change coincidentally with corporate taxes.

Denoting T_c^n as the date of the nth corporate tax change in country c, the dummy $D_{ct}^j = 1[t = T_c^n + j]$: indicates that country c experiences a tax change j periods away. The main regression is as follows:

$$Y_{ckpt} = \sum_{j=\underline{\omega}}^{\overline{\omega}} \beta_j^{\text{side}} D_{ct}^j \text{tax rate}_{ct} + \gamma \, \textbf{controls}_{\boldsymbol{ct}} + \text{FE}_{ckp} + \text{FE}_{kpt} + \text{FE}_{ckpf} + \epsilon_{ckpt} \tag{1.18}$$

where the country in which the tax changes is denoted as c and the partner country for a given bilateral trade flow is denoted as k. The subscript $f \in \{m, q\}$ denotes the frequency of the data, either months or quarters. When studying tax changes in the importing country, $c = i$ and $k = e$, and conversely when focusing on tax changes in the exporting country. The exponent on β_j indicates which side is studied, i.e. $\text{side} = \{\textit{importer} \text{ if } c = i, \textit{exporter} \text{ if } c = e\}$. Focusing on one side at a time allows the inclusion of a country-time fixed effect for the partner country.

The coefficients of interest are the sequence of β_j. Importantly, this specification allows to check for parallel pre-trends: the β_js for the periods leading to a tax change (i.e. when j is negative) must not be significantly different from zero. If this were the case, it would indicate that the treated country — i.e. that which experiences a tax change — is on a different trend before the tax change takes place.[40]

---

[40] For papers using a similar setup in other contexts, see Fuest et al. (2018); Suárez Serrato and Zidar (2016); Drechsler et al. (2017). For recent studies on the assumptions and data requirements needed to run an event study, see Schmidheiny and Siegloch (2019); Abraham and Sun (2018); Borusyak and Jaravel (2017).

1.4.3 Threats to identification

There are two main threats to properly identifying the effects of changes in tax and tariff rates on the reporting gaps in the panel regressions.

Reverse causality If tariffs and/or corporate taxes were set in response to the level of evasion, the regressions above would suffer from endogeneity. It is in principle possible that tariffs are set taking evasion risks into account. For instance, tariffs could be lower for products on which duties are known to be easily and often evaded. The inclusion of importer-exporter-product fixed effects should neutralize this worry, unless evasion becomes apparent over time for a given importer-exporter-product observation, and tariffs are changed accordingly. Second, most countries in the sample are part of the WTO, wherein tariffs cannot be freely set by individual countries, but are the fruit of multilateral negotiations — as is the case for any trade agreement. This limits the freedom a country may have to set tariffs in response to evasion observed domestically.

Reverse causality seems unlikely in the case of corporate taxes. First, the rate of taxation is set in complex political negotiations, usually in the yearly budget laws. The rate is often changed to meet the needs of governments and in response to macroeconomic shocks to revenues or expenses — often correlated to GDP levels that I control for in regressions. Anecdotally, I have never seen tax evasion as a reason for a tax change whilst reading through the laws amending the tax rates. Tørsløv et al. (2018) note that the marked downward trend in CIT rates in OECD countries appears to be driven by tax competition between countries striving to attract capital and profits from abroad. In both the cases of tariffs and corporate taxes, my estimates of evasion would be biased downwards if rates were lowered in response to high evasion levels.

Omitted variable bias Any variable correlated with both taxes and the reporting gap will result in an omitted variable bias (OVB) unless it is included in the regression. Using the terminology introduced in section 1.4.1, the components of χ_{eipt} must either be controlled for or uncorrelated with taxes.

Freight and insurance costs are unlikely to be systematically related to tariffs or taxes. Furthermore, controlling for country pair, product and year fixed effects is arguable sufficient to account for trade costs, which may be related to the distance between countries, characteristics of the product, or the cost of fuel. As a robustness check, the correlation between tax rates and CIF-FOB margins is explored for a subset of products where data on the latter is available.[41]

Measurement errors — in a broad sense, including methodological differences in data collection and treatment — could be related to tariffs and corporate taxes. Regarding tariffs, it is known that imports are generally more accurately recorded as they form the tax base for custom duties (Javorcik and Narciso, 2017, p. 63). If a higher tariff warranted more attention from customs and thus more accurate recording, evasion would be lower for these products (i.e. the estimated effect of tariffs on evasion would be a lower bound). Country-time fixed effects account for any changes in data collection methods at the country level. When studying corporate tax evasion, country-time fixed effects have to be dropped for the country in which taxes change. If measurement errors change together with the tax rates, there will be an OVB. The event study intends to alleviate this concern.[42] Furthermore, a series of model predictions will be tested to lend credibility to a causal interpretation of the results.

Changes in the tax base Changes in the tax rate are sometimes accompanied by changes in the tax base, which may impact the appetite for evasion. The tax base definition can change at the same time as the rate of tax. (Kawano and Slemrod, 2016) document that the tax base definition in more likely to change when the rate is also changed in OECD countries. See also Serrato and Zidar (2018) on the link between CIT revenues, rate and base. Several sources suggest that in the OECD, the general decrease in CIT rates was accompanied by a broadening of the tax base (Clausing, 2007; Becker and Fuest, 2011).[43] If both rate and base change simultaneously, it is difficult to attribute the estimated change in evasion to the rate alone. As a robustness check, changes in the tax base are controlled for in a subset of the OECD sample where the data is available. The results can be found in section 1.5.4.

The next two sections detail the results of the paper, starting with estimates of corporate income tax evasion, and then discussing the evidence of tariff evasion. In

---

[41] Miao and Fortanier (2017) provide a review of the literature attempting to measure CIF-FOB margins, and a new measure thereof. One approach is to use explicit data on insurance and freight costs. Another approach uses bilateral asymmetries — i.e. the reporting gaps — as a proxy for the CIF-FOR margin. They collect explicit CIF-FOB margins data where available, estimate a gravity-type equation to predict them, and use that estimated model to infer the margins for countries and products where the explicit data is not available.

[42] Country-specific time trends are also included in some specifications of the panel regression.

[43] Despite the decline in CIT rates, the importance of CIT revenues relative to total tax revenues remained constant at around 8-10% on average in the OECD since the early 1990s.

general, the relation between tax rates and reporting gaps must not be considered as causal with certainty. In the following sections, multiple empirical exercises exploiting heterogeneity in the link between gaps and taxes in an effort to facilitate causal interpretation of the results.

1.5 Corporate Income Tax evasion

In this section, the results of the investigation of corporate tax evasion are detailed. First, the focus is on testing the basic predictions of the model, i.e. how do the reporting gaps react to changes in the CIT rates of the importing and the exporting countries. Second, model-motivated predictions from section 1.2 based on interactions between different factors are tested empirically to refine the analysis and buttress causal interpretation. Third, additional robustness checks are performed, namely controlling for changes in the tax base, and replicating the baseline results using data on trade in services. Last, seasonality features of the reporting gaps at a monthly frequency are linked to evasion motives, and reactions to tax changes are studied in an event study framework.

1.5.1 Baseline results

OECD sample The importer-exporter-year fixed effects must be dropped from the regression to allow corporate income tax rates as regressors. Identification is therefore more challenging and the risk of OVB increases. Coefficients should be interpreted as correlations at this stage. Table 1.4 reports how the reporting gaps change when the corporate income tax in the importing and exporting countries change — corresponding to regression equation (1.15). In these regressions, both the importer's and the exporter's tax rates are included simultaneously, which restricts the sample to within-OECD trade as data on the CIT rate must be available for both countries. All regressions include the set of controls listed in section 1.4. Columns (1) and (2) are estimated without country-specific time trends, whereas columns (3) and (4) include these. As a reminder, the model predicts a negative effect of the CIT rate faced by exporters on the reporting gaps — as exporters under-report exports when the CIT increases, cf. prediction 2 — whereas the effect of the CIT rate faced by importers has an ambiguous effect and depends on whether condition 1.9 holds. The results suggest that reporting gaps do not change when the corporate income tax rate in the importing

TABLE 1.4: CORPORATE INCOME TAX RATES AND REPORTING GAPS: BOTH SIDES

	(1) Value gap, log	(2) Quantity gap, log	(3) Value gap, log	(4) Quantity gap, log
CIT Exporter	-0.00205***	-0.00140	-0.00115	-0.00230**
	(-2.94)	(-1.56)	(-1.63)	(-2.49)
CIT Importer	0.000839	0.000783	0.000302	-0.000543
	(1.08)	(0.88)	(0.37)	(-0.57)
Tariff	-0.000970	-0.00312***	-0.000933*	-0.00210**
	(-1.37)	(-3.00)	(-1.85)	(-2.20)
Importer-exporter-product FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Importer-specific time trend			✓	✓
Exporter-specific time trend			✓	✓
Controls	✓	✓	✓	✓
Adjusted R^2	0.319	0.314	0.321	0.316
Within R^2	0.000	0.001	0.000	0.000
Observations	27,002,571	23,706,919	27,002,571	23,706,919

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: OECD countries. CIT refers to the central CIT.

country changes. On the exporter's side, a one percentage point increase in the CIT rate is associated with a decrease in the gap of 0.1% to 0.2%, consistent with exporters under-reporting exports to lower their taxable profits.

Investigating CIT evasion on each side separately — i.e. estimating regressions (1.16) and (1.17) — presents two distinct advantages. First, it allows to use a larger sample as the partner countries do not need be in the OECD.[44] Second, it permits the inclusion of partner-time fixed effects that control for any partner-specific variation in the reporting gaps over time. This includes the corporate income tax rate in the partner country. The drawback is that the samples for each side slightly differ. Table 1.5 details the results. The negative association between the gaps and the exporter's CIT is confirmed. There is still no discernible relation between the gaps and the importer's CIT. The results remain largely unchanged if country-specific time trends are excluded from the regressions (cf. table 1.A.4), for aggregated trade at the HS 2-digit level (cf. table 1.A.5), and for regressions in first differences (cf. table 1.A.6).[45] Furthermore, the results remain qualitatively unchanged when the regressions are estimated (i) using a more balanced sample (cf. table 1.A.7); (ii) using the alternative measure of the gaps in expression (1.12) (cf. table 1.A.8); (iii) using the log of net-of-tax rate instead of the tax rate itself as regressor, although the interpretation of the coefficients change, which

---

[44] In terms of terminology, the *partner* country refers to the country on the other side of the reporting gap, i.e. where CIT evasion is not being scrutinized in a given regression.

[45] Note that trade can only be aggregated in terms of value, as the quantity units vary across products (e.g. tons and meters cannot be added)

TABLE 1.5: CORPORATE INCOME TAX RATES AND REPORTING GAPS: EACH SIDE SEPARATELY

	(1) Value gap, log	(2) Quantity gap, log	(3) Value gap, log	(4) Quantity gap, log
CIT Exporter	-0.00150***	-0.00232***		
	(-3.62)	(-4.16)		
CIT Importer			0.000582	-0.000348
			(0.92)	(-0.51)
Tariff	0.000963***	-0.00135***	-0.000650	-0.00107
	(4.10)	(-3.19)	(-1.59)	(-1.61)
Importer-export-product FE	✓	✓	✓	✓
Importer-year FE	✓	✓		
Exporter-year FE			✓	✓
Exporter-specific time trend	✓	✓		
Importer-specific time trend			✓	✓
Controls	✓	✓	✓	✓
Adjusted R^2	0.326	0.322	0.335	0.329
Within R^2	0.000	0.000	0.000	0.000
Observations	44,623,145	38,666,718	33,866,668	29,732,923

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: OECD countries. CIT refers to the central CIT.

makes quantitative comparisons to the baseline results difficult (cf table 1.A.9).[46] The estimated elasticity of reported sales with respect to the net-of-tax rate is 0.1. This elasticity is easier to compare to other studies than the baseline semi-elasticity. Devereux et al. (2014) find an elasticity of taxable income with respect to the marginal tax rate of 0.14 to 0.18 for firms with profits around £300,000 in the UK. Gruber and Rauh (2007) estimate this elasticity to be 0.2 for US public companies, based on accounting data. These objects are different since the elasticity of taxable income also captures changes in real economic activity due to the tax change, whereas only evasion is meant to be captured in the present study.

World sample Using the combined CIT allows for a larger sample, with up to 170 countries. The disadvantage is that the combined CIT includes regional taxes than may not affect all traders. All the regressions presented above are re-estimated using this larger sample. The main results remain largely unchanged and can be found in table 1.A.10 and 1.A.11 (with and without country-specific time trends, respectively). Note however that OECD countries account for a significant part of the observations in the world sample, so non-OECD countries may have limited impact on the results, perhaps contributing to the stability of the regression coefficients.

---

[46] To balance the sample, observations appearing in less that 80% of the periods since their first appearance in the panel are dropped.

TABLE 1.6: INTERACTION OF CORPORATE TAX AND TARIFFS RATES

	(1) Value gap, log	(2) Quantity gap, log	(3) Value gap, log	(4) Quantity gap, log
CIT Importer	0.000853	0.000115	0.000951	0.00102
	(1.34)	(0.17)	(1.58)	(1.48)
Tariff	0.00483***	0.00805***	0.00784***	0.0105***
	(3.14)	(3.58)	(4.59)	(4.76)
CIT Importer × Tariff	-0.000187***	-0.000311***	-0.000291***	-0.000438***
	(-3.47)	(-3.68)	(-5.04)	(-5.29)
Importer-exporter-product FE	✓	✓	✓	✓
Exporter-year FE	✓	✓	✓	✓
Importer-specific time trend	✓	✓		
Controls	✓	✓	✓	✓
Adjusted R^2	0.335	0.329	0.334	0.328
Within R^2	0.000	0.000	0.000	0.000
Observations	33,866,668	29,732,923	33,866,668	29,732,923

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: OECD countries. CIT refers to the central CIT. Columns (1) and (2) include an importer-specific time trend, whereas columns (3) and (4) do not.

Summary Higher corporate income taxes in the exporting country are associated with lower reporting gaps, consistent with prediction 2: exporters lower their reported exports to reduce taxable profits. However, there does not seem to be strong co-movements between gaps and CIT rate in the importing country. However, the model does not deliver a clear-cut prediction regarding this relationship, which depends on whether equation (1.9) holds or not, i.e. whether importers find it optimal to over- or under-report imports upon an increase in the CIT rate. The next section focuses on exploring the correlations between gaps and importers' CIT rate under different scenarios in which condition (1.9) is more or less likely to hold.

1.5.2 Further tests of model-motivated predictions

In order to facilitate causal interpretation of the results above, a series of non-trivial predictions from the model are tested in this section. They mostly rely on variation across products or countries in dimensions that make condition (1.9) more or less likely to hold — i.e. that make the importer under- or over-report imports when the CIT rate changes.

Interaction between tariffs and CIT As seen in the model, importers have different incentives when it comes to evading tariffs or corporate income tax. On the one hand, imports should be under-reported to avoid tariffs. On the other hand, imports could be over-reported to lower taxable profits and evade CIT. Furthermore, the level of each tax affects the incentives to evade the other tax. For instance, high tariffs increase the incentive to over-report imports, as each dollar over-reported lowers taxable profits by a factor of $\phi(1+t)$. This is captured in condition (1.9): high tariffs make it less likely to hold and make the firm more likely to over-report imports. Seen from another angle, a high CIT rate lowers the incentive to evade tariffs by under-reporting imports, as savings from tariff evasion increase profits that are heavily taxed. To test this hypothesis, regression (1.17) is estimated adding an interaction between the CIT rate in the importing country and tariffs. The results can be found in table 1.6. The interaction term has the expected sign: higher tariffs result in a more negative association between the importer's CIT and the reporting gaps. Put differently, a higher CIT rate lowers the incentive to evade tariffs. The coefficient on the CIT is positive (but not significant) also consistent with condition (1.9) holding when tariffs are zero. The fact that this interaction — which is predicted by the model — finds support in the data is one step towards a more causal interpretation of the results.

Within-EU trade Within-EU trade offers an environment in which there are no tariffs, which mitigates concerns of interactions between incentives to evade tariffs and corporate taxes — see above. Furthermore, as explained in section 1.3.2.2, there are no customs within the EU, which effectively removes one way tax authorities can check for fraud. However, since trade statistics are gathered via VAT returns and Intrastat declarations, the taxman has access to reports of imports and exports on the same form, which arguably makes detection of inconsistencies between imports and exports of a given firm more likely.[47] This institutional feature is interpreted as a high value of $|C_{12}|$ in the cost function. Combined with zero tariffs, condition (1.9) is more likely to hold, i.e. importers more likely to under-report imports upon an increase in their CIT rate.

The results restricting the sample to within-EU trade can be found in table 1.7 — and in table 1.A.12 for regressions without country-specific time trends. The relevant Intrastat thresholds in each country and the standard VAT rate are added to the regressions as controls in all regressions based on within-EU trade flows. The reporting gaps are negatively associated with the CIT rate in the exporting country, consistent with model prediction 2 and with previous findings. Furthermore and unlike in the full

---

[47] For instance, the HMRC in the UK uses a computer program that mines data across a variety of sources and looks for discrepancies. See here (last accessed 02/08/2019).

TABLE 1.7: CORPORATE INCOME TAX RATES AND REPORTING GAPS: WITHIN-EU TRADE

	(1) Value gap, log	(2) Quantity gap, log	(3) Value gap, log	(4) Quantity gap, log
CIT Exporter	-0.00289***	-0.00387***		
	(-4.63)	(-4.63)		
Log(Intrastat threshold), Exporter	-0.0380***	-0.0459***		
	(-4.79)	(-4.60)		
Standard VAT rate, Exporter	0.00226	0.00390		
	(1.00)	(1.43)		
CIT Importer			0.00269***	0.00401***
			(3.16)	(4.81)
Log(Intrastat threshold), Importer			0.0866***	0.0925***
			(9.54)	(9.80)
Standard VAT rate, Importer			-0.000434	-0.000829
			(-0.16)	(-0.29)
Importer-export-product FE	✓	✓	✓	✓
Importer-year FE	✓	✓		
Exporter-year FE			✓	✓
Exporter-specific time trend	✓	✓		
Importer-specific time trend			✓	✓
Controls	✓	✓	✓	✓
Adjusted R^2	0.319	0.310	0.321	0.310
Within R^2	0.000	0.000	0.000	0.000
Observations	16,899,957	15,167,030	15,891,282	14,271,876

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: EU countries. CIT refers to the central CIT.

OECD sample, the gaps are positively associated with the CIT rate in the importing country, consistent with model prediction 3 when condition (1.9) holds. This suggests that importers under-report the value of their imports when the CIT rate increases. In the context of the model, this happens as firms under-report their sales and their costs simultaneously in order to avoid large discrepancies in their reported gross margins, which would increase the likelihood of being caught. This result suggests a behaviour similar to that found by Carrillo et al. (2017) and Slemrod et al. (2017), where both sales and costs are adjusted together in the same direction when misreported revenues are more likely to be discovered by tax authorities.[48]

---

[48] Note that changes in the Intrastat reporting thresholds are expected to affect the reporting gaps. An increase in a threshold means that less firms are subject to reporting for statistical purposes. Specifically, an increase in the imports thresholds should therefore result in lower imports and higher gaps, and an increase in the exports thresholds should imply lower exports and lower gaps. This is confirmed by the data. In the specifications with time trends, VAT rates do not affect the reporting gaps. This is contrary to findings by Gradeva (2014) and Frunza (2016) in studies of VAT fraud. These studies relate the VAT rate to the reporting gaps in order to find evidence of the Missing Trader Intra-Community (MTIC) fraud scheme. This scheme basically consists in importing goods — on which VAT is not payable at the border, since there are no customs, but at the time when the firm files its VAT return — selling them to a domestic purchaser and charging VAT on that sale, and disappearing with the VAT amount charged before having paid VAT on the imports. European Commission (2018b) reports estimates of VAT revenue losses ranging from EUR 20 to EUR 100 billion a year from MTIC fraud. Note that in tables 1.7 and 1.A.12, the regressions are estimated on all products, including products that are subject to reduced VAT rates — typically a small subset of all products. Finally, I also estimate the regressions dropping all products susceptible to be subject to the Domestic Reverse Charge Mechanism (see Bussy, 2020) and the results remain stable.

A natural question following the results above is: in response to an increase in the CIT rate, do importers over-report imports if one excludes within-EU trade from the sample (where importers clearly under-report imports)? Regression (1.17) is estimated using the whole OECD sample, augmented with an interaction of the CIT rate in the importing country and a dummy indicating that both countries are EU members (i.e. indicating within-EU trade). The results are displayed in table 1.A.13: the reporting gap react negatively to an increase in the CIT rate for trade flows that are not between EU members, and positively for within-EU trade. In other words, the absence of relationship between reporting gaps and importers' CIT found in the baseline results (cf. table 1.5) masked the fact that within the EU importers appear to under-report, and in the rest of the sample importers seem to over-report (in response to an increase in the CIT rate).

Capital and accounting rules Not all goods can immediately be accounted as costs against tax by firms. Specifically, capital goods must generally be first accounted as assets, and then gradually turned into costs over the lifetime of the good according to a depreciation schedule (Devereux et al., 2002). Depreciation allowances are one of the tools that can be used by tax authorities to modify the tax base and the generosity of the tax system (Kawano and Slemrod, 2016; Becker and Fuest, 2011). Even if ultimately the whole value can be depreciated — which is not always the case — spreading the process over years results in a net present value of allowances inferior to the original value of the good. In the model, this is captured in reduced form by parameter ϕ. When it comes to optimal reporting, the model predicts that importers are less likely to over-report imports (in response to a CIT rate increase) when ϕ is low, i.e. when the value cannot be fully accounted as costs upon purchase, as captured by condition (1.9).

A comparison of how the gaps react to changes in the CIT rate in the importing country for capital goods relative to all other goods is provided. This is done by interacting the tax rate with a dummy indicating whether the product is a capital good in regression (1.17).[49] The results can be found in table 1.8. Strikingly, capital goods appear to be under-reported by importers as the CIT rate increases, as expected given the above arguments. Reported imports of other goods appear not to react to changes

---

[49] The classification of products into different goods categories is provided by WITS and can be found here (last accessed 16/10/2019).

TABLE 1.8: CORPORATE INCOME TAX RATES AND REPORTING GAPS: CAPITAL GOODS

	(1) Value gap, log	(2) Quantity gap, log
CIT Importer	0.000217	-0.000875
	(0.34)	(-1.17)
Capital goods × CIT Importer	0.00153***	0.00272**
	(3.39)	(2.01)
Importer-export-product FE	✓	✓
Importer-year FE		
Exporter-year FE	✓	✓
Exporter-specific time trend		
Importer-specific time trend	✓	✓
Controls	✓	✓
Adjusted R^2	0.335	0.329
Within R^2	0.000	0.000
Observations	33,866,668	29,732,923

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: OECD countries. CIT refers to the central CIT. Note that the dummy indicating a capital good cannot be included in the regression, as it would be absorbed by the product fixed effect.

in the CIT, as previously found in the OECD sample.

Fight against evasion The extent of evasion is likely to be lower when governments actively fight against it. This is the argument that will be explored here. The first question to address is: how should the estimated effects of CIT rates on the reporting gaps be affected by a crackdown on evasion? From the model, higher effort to fight evasion can be interpreted as a higher value of parameter b in front of the cost function, as well as a higher value of $|C_{12}|$. If these two parameters change simultaneously, the ultimate effects on the relationships between gaps and tax rates can be ambiguous. First, consider the exporters' incentives. Increases in $|C_{12}|$ and b both induce exporters to under-report exports to a lesser extent since (i) a higher value of b scales up the evasion cost and (ii) a higher value of $|C_{12}|$ implies less deviations in terms of reported gross margin, which forces the exporter to under-report less in order to minimize deviations in that dimension. Therefore, a higher fight intensity against evasion unambiguously leads the exporter to under-report less: β^{exp} is expected to be less negative. Second, consider the importers' incentives. A higher value of $|C_{12}|$ will make reported imports more aligned with reported exports, which leads importers to over-report less — and can even lead them to switch from over-reporting to under-reporting imports (condition (1.9) is more likely to hold for high values of $|C_{12}|$). This would imply a higher value of β^{imp}. However, an increase in b induces importers to lower the extent to which they misreport in either direction: they under-report less if they under-report, and they over-report less if they over-report. Therefore, the effect of b on optimal reported imports depends on whether condition (1.9) holds or not. If it does and importers under-report, the effects of b and C_{12} counteract each other, yielding ambiguous predictions regarding the effects of evasion-reducing policies on β^{imp}.

In order to test the mechanisms presented above, a measure of governments' effort to curb corporate income tax evasion is obtained from Kawano and Slemrod (2016). They document changes to the tax base in OECD countries between 1980 and 2004 along twelve dimensions, one of which consists in policies that target evasion or avoidance by companies.[50] Specifically, they provide a variable that can take values -1, 0 or 1 for each of the twelve dimensions, 1 meaning that a change in that dimension that widens the tax base has occurred, and -1 that a change that shrinks the tax base has occurred. These variables are available annually for all countries in the sample. Focusing for now solely on policy changes targeting tax evasion, the cumulative changes over time in that dimension in each country is used as a proxy for the level of fight against evasion. Descriptive statistics of this variable can be found in table 1.A.15. Some countries have implemented many policies against corporate income tax evasion, such as the UK, Australia and Italy, with up to 12 such instances.

Evidence that evasion appears less prevalent when governments make an active effort to curb it is provided by including an interaction of the cumulative change in evasion crackdown as defined above and the tax rates in regressions (1.16) and (1.17). The results can be found in table 1.9. Columns (1) and (3) do not include the interaction term, whilst columns (2) and (4) do. Consistent with model the predictions discussed above, the association between gaps and CIT rate in the exporting country is less negative when governments fight harder against evasion. There is no discernible effect on the importer side, which mirrors the ambiguity of the model predictions discussed above.

1.5.3 Evasion amounts

The estimated reaction of reporting gaps to changes in CIT rates is interpreted as evidence of evasion. However, international trade is only used as a lens through which evasion is detected. Estimating the *levels* of CIT evasion is thus more complicated than

---

[50] A list of these dimensions can be found in table 1.A.14.

TABLE 1.9: CORPORATE INCOME TAX RATES AND REPORTING GAPS: CRACKDOWN ON EVASION

	Dep. variable: Value gap, log			
	(1)	(2)	(3)	(4)
CIT Exporter	-0.000987**	-0.00455***		
	(-2.06)	(-4.77)		
Cum. Δ Ev. fight, Exp.		-0.0393***		
		(-6.55)		
CIT Exporter × Cum. Δ Ev. fight, Exp.		0.000874***		
		(5.69)		
CIT Importer			0.00176**	0.00190
			(2.23)	(1.33)
Cum. Δ Ev. fight, Imp.				0.0164*
				(1.95)
CIT Importer × Cum. Δ Ev. fight, Imp.				-0.000133
				(-0.62)
Importer-export-product FE	✓	✓	✓	✓
Importer-year FE	✓	✓		
Exporter-year FE			✓	✓
Exporter-specific time trend	✓	✓		
Importer-specific time trend			✓	✓
Controls	✓	✓	✓	✓
Adjusted R^2	0.391	0.391	0.409	0.409
Within R^2	0.000	0.000	0.000	0.000
Observations	13,441,475	13,441,475	10,502,506	10,502,506

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: OECD countries, 1988-2005. CIT refers to the central CIT. The variable *Cum. Δ Ev. fight* refers to the cumulative change in policies aiming at curbing CIT evasion as defined by Kawano and Slemrod (2016).

in the case of tariffs, where the tax base is trade itself. Specifically, it is not clear that the extent to which firms manipulate their reported exports and imports and their domestic sales and purchases is the same. For instance, it could be easier to manipulate transactions that involve non-domestic firms if it is harder for the tax authorities to obtain information on foreign firms that are not directly in their jurisdiction, which seems a plausible assumption. A tentative figure of the amounts evaded for a country c in year t could be calculated as follows:

$$\text{CIT evaded}_{ct} = -\left[F^D_{ct}\epsilon^D_F + F^A_{ct}\epsilon^A_F + M^D_{ct}\epsilon^D_M + M^A_{ct}\epsilon^A_M\right] \times \text{CIT}_{ct} \tag{1.19}$$

where F^D_{ct} is domestic sales, F^A_{ct} denotes exports, M^D_{ct} is intermediate consumption from domestic sources and M^A_{ct} denotes imports (A stands for *abroad* and D for *domestic*). ϵ^j_k is the semi-elasticity of reported amounts with respect to the CIT rate, for $k \in \{F, M\}$ and $j \in \{D, A\}$. From the estimation of equations (1.16) and (1.17), $\epsilon^A_F = \hat{\beta}^{exp}$ and $\epsilon^A_M = \hat{\beta}^{imp}$. However, the values of ϵ^D_k, i.e. the semi-elasticities of reported sales and purchases for domestic transactions, are unknown. For the purpose of estimating the amount of CIT evasion, basing the calculations solely on imports and exports is likely to give biased estimates as the volumes of these flows do not reflect those of total purchases and sales. Consider the extreme example where a country imports all its intermediate goods and does not export at all: depending on the values of $\hat{\beta}^{exp}$ and $\hat{\beta}^{imp}$, negative estimates of evasion amounts could arise. More generally, countries with large trade imbalances will significantly weight on the resulting estimated CIT evasion amounts. A tentative estimate of evasion levels stemming from misrepresenting international trade only, which should be taken with extreme caution, is calculated by taking the average across countries of each elements of expression (1.19) and setting ϵ_k^D for $k \in \{F, M\}$ to zero.[51] Taking the average across countries reduces the issue of imbalances since trade is more balanced at the OECD level. The result is a time series of average evasion for the OECD as a whole. On average, evasion calculated in this way represents around 16% of actual corporate income tax revenues. This estimate is in the same ballpark as estimates of CIT evasion from tax gap calculations and random audits in different countries, which approximately range from 9% to 30% (European Commission, 2018a). Yet, it represents solely evasion and not legal avoidance, as opposed to some measures of the tax gap. This estimate also suffers from the fact that evasion is assumed to be constant at every level of CIT rates, which is unlikely to be true.

1.5.4 Additional Robustness checks

1.5.4.1 Changes in the tax base

One well known phenomenon in the literature on corporate income tax is that in many high-income countries, changes in the statutory tax rate are often accompanied by changes in the tax base (Kawano and Slemrod, 2016; Becker and Fuest, 2011; Devereux et al., 2002). This may be an issue for identification if changes in the tax base modify firms' incentives to evade. For example, if the tax rate decreases and the tax base widens simultaneously — a pattern that seems prevalent among OECD countries — one would predict evasion to lessen as a result of the decrease in the CIT rate, but

---

[51] Setting $\epsilon_F^D = \epsilon_M^D = 0$ does not necessarily mean that the resulting evasion level estimates are lower-bounds for the overall evasion levels (i.e. that include evasion from misreported domestic transactions) if firms were to misreport in such a way that under-reporting international trade is compensated by over-reporting domestic transactions, or vice-versa. Although it is not intuitive to imagine such a case, it cannot be categorically excluded.

perhaps to increase as a result of the widening of the tax base.[52] Using the data by Kawano and Slemrod (2016) on changes in the corporate income tax base introduced in section 1.5.2, a measure of cumulative changes in the tax base is calculated as the cumulative sum of changes across all dimensions on which the authors provide data. Table 1.A.14 contains a list of the dimensions and summary statistics of the cumulative changes in the tax base can be found in table 1.A.15. The UK is the country that most widened its tax base (mostly through policies to curb evasion), whereas Luxembourg shrank its base most markedly (through investment credit and generous depreciation allowances).

This measure is added as an additional control in regressions (1.16) and (1.17). Of interest is whether the coefficients on the CIT rates remain stable after controlling for changes in the tax base. Table 1.10 contains the results for OECD countries over years 1988-2005. Columns (1) and (3) display the results without controlling for changes in the tax base, whereas regressions in columns (2) and (4) do control for base changes. The coefficients are very stable, although the coefficient on the importer's CIT is slightly smaller once changes in the tax base are controlled for. Although this robustness check can only be performed on a subset of the data, the results suggest that omitting changes in the tax base definition does not result in a substantial bias. Note that changes in the tax base themselves could also impacts the reporting gaps in principle. It is however difficult to interpret these coefficients since this measure is a composite of changes in 12 distinct dimensions of the tax base.

One could also use measures of effective average tax rates (Devereux and Griffith, 1998, 2003) instead of statutory rates in regressions. Effective average tax rates (EATR) are synthetic measures of the tax payment of a firm on an investment that yields a positive return based on the comparison of the net present values of the investment in the presence and absence of tax. EATR take into account depreciation, source of financing and country-specific tax features affecting the final tax bill. The advantage is that EATR may be more representative of the tax payment faced by firms, yet the complexity of the calculations and of the underlying assumptions render interpretation of regression coefficients challenging. I estimate the baseline regressions using EATR from Spengel et al. (2019) for European Member States for years 1998-2017.[53]

---

[52] If this were true, omitting to control for changes in the tax base would result in underestimating evasion, since a decrease in the rate lowers evasion, yet not as much when the base simultaneously widens.

[53] Specifically, I use the overall mean of the EATR at the corporate level. Data availability defines the

TABLE 1.10: CORPORATE INCOME TAX RATES AND REPORTING GAPS: CHANGES IN TAX BASE

	Dep. variable: Value gap, in log			
	(1)	(2)	(3)	(4)
CIT Exporter	-0.000987**	-0.00100**		
	(-2.06)	(-2.02)		
Cum. Δ base, Exp.		-0.000209		
		(-0.10)		
CIT Importer			0.00176**	0.00168**
			(2.23)	(2.15)
Cum. Δ base, Imp.				-0.00364**
				(-2.11)
Importer-export-product FE	✓	✓	✓	✓
Importer-year FE	✓	✓		
Exporter-year FE			✓	✓
Exporter-specific time trend	✓	✓		
Importer-specific time trend			✓	✓
Controls	✓	✓	✓	✓
Adjusted R^2	0.391	0.391	0.409	0.409
Within R^2	0.000	0.000	0.000	0.000
Observations	13,441,475	13,441,475	10,502,506	10,502,506

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: OECD countries, 1988-2005. CIT refers to the central CIT. The variable *Cum. Δ base* refers to the cumulative change in the tax base as defined by Kawano and Slemrod (2016).

The magnitudes of the coefficients double relative to regressions with the statutory CIT rate, as can be seen by comparing columns (1)-(2), and (3)-(4) in table 1.A.16. Since the difference between EATR and statutory rates could stem from so many factors, it is difficult to confidently explain this difference.

1.5.4.2 Trade in services

Trade in services is generally less important than trade in goods, accounting for only 14% of GDP on average in 2018 (both among OECD countries and in the world) against around 45% for trade in goods.[54] It appears however to be a fertile ground for evasion and tax avoidance for two main reasons. First, it does not involve a physical movement of merchandise across borders. Second, services are differentiated goods for which reference prices are generally not readily available.[55] For instance, it has been documented in the literature that trade in services is used by multinational firms for shifting profits to low-tax countries (Hebous and Johannesen, 2015; Tørsløv et al., 2018). Trade in services also appears to be used for money laundering purposes for

---

sub-sample.

[54] Source: World Bank series *TG.VAL.TOTL.GD.ZS* and *BG.GSR.NFSV.GD.ZS*.

[55] This insight was gained speaking to compliance professionals in trade finance. Transactions involving goods and services associated with the goods (e.g. maintenance or repair services) are also considered as risky from a compliance point of view since determining what a reasonable price for the bundle is difficult.

TABLE 1.11: DESCRIPTIVE STATISTICS: TRADE IN SERVICES

	N	Mean	Std	Min	25%	50%	75%	Max
Value gap, in log	195,946	0.02	1.68	-11.69	-0.83	0.01	0.86	12.21
Value gap, robust	195,947	0.01	1.06	-2.00	-0.78	0.01	0.81	2.00

Note: descriptive statistics of the reporting gaps in terms of value in the services trade data.

the same reasons (Sullivan and Smith, 2012). Given the observations above, it is possible that CIT evasion may also be detectable in services trade flows, and it may be of higher magnitude than in the case of trade of physical goods.

Data Data on international trade in services is of lesser quality relative to data on trade in goods, mainly due to the intangible nature of services and the fact that tariffs levied on trade in goods induces countries to keep accurate records for tax collection purposes (Francois and Pindyuk, 2013). Comtrade collects data on trade in services from national statistical agencies. The data comes from Balance of Payment statistics and trade flows are classified according to the Extended Balance of Payments Services Classifications (EBOPS), a coarse classification involving 12 main categories and several finer aggregation levels - down to about 60 categories (UNSD, 2002, pp. 30-32). Given the difficulty of gathering accurate partner-specific data on trade flows of services, the data is often missing, and is skewed towards the main partners of each reporting country. Furthermore, there is suggestive evidence that some countries may use bilateral statistics (i.e. the reports of the partner country) to infer trade flows for the production of their own data, which is a serious issue in the present case since the bilateral asymmetry is precisely the dependent variable in regressions.[56] As in the case of trade in goods, bilateral asymmetries in services' trade flows are known and discussed among data providers.[57] The causes are generally similar to these that apply to trade in goods as listed in section 1.3.2.4.

The data is available over the years 2000-2017 for 43 countries — mostly high-income. The main analysis is restricted to OECD and EU countries. The sample is unbalanced, with many observations only observable since 2004 and up to 2016. Summary statistics and histograms of the reporting gaps — computed using the same formulas as the main analysis — can be found in table 1.11 and figure 1.B.14, respectively. The descriptive statistics of the gaps are very similar to that in the case of trade

---

[56] *UNSD* (2019a), section on *Bilateral Trade Asymmetries in Trade in Services*. The world bank also provides a database on trade in services where missing flows are explicitly estimated using bilateral statistics (WB *Trade In Services Database*).

[57] See for instance reports by Comtrade, the UK ONS and the BEA (last accessed 09/09/2019).

in goods. However, the histograms reveal much fatter tails — especially in the robust measure of the gaps — consistent with lower-quality data. Considering the quality challenges presented by the data, the results in this section ought to be taken with caution. One advantage however is that the data collection processes for services and goods differ markedly. Finding similar results to those in section 1.5.1 would partially alleviate concerns that measurement errors inherent to merchandise trade data bias the estimates.

Results Regressions (1.15), (1.16) and (1.17) are estimated using services trade data. The results when including the CIT rates of both sides in the regression can be found in table 1.12. The results are qualitatively almost identical to the goods trade benchmark: an increase in the CIT of the exporter is associated with lower gaps. An increase in the CIT of the importer does not seem to impact the gaps in the OECD sample, yet is positively associated with the gaps in the within-EU sample. Quantitatively, the coefficients are larger — up to one order of magnitude larger in some cases. This could be because of the quality of the data, or because evading using services is much easier. The same regressions are estimated focusing on one side at a time and the results can be found in table 1.A.17. As in the case of trade in goods, these regressions are considered more robust as they maximize the number of observations and allow to control for partner-year fixed effects. The results are very similar, although the magnitudes of the coefficients are somewhat smaller. The results remain unchanged using the alternative definition of the reporting gaps and estimating the regressions on a more balanced sample. However, including partner-specific time trends renders all coefficients statistically insignificant. This is perhaps due to the short time horizon and the fact that relatively few tax changes take place during that time.

Despite the data challenges, these results are comforting. First, all coefficients are qualitatively similar to these estimated based on trade in goods. Second, the magnitudes are larger, consistent with the notion that services are relatively easier to misreport and manipulate for fraudulent purposes.

1.5.5 Dynamics at the monthly frequency

This section is based on the within-EU trade sample, available at a monthly frequency. First, a striking pattern of seasonality of the reporting gaps is reported, which suggests that profits are shifted from a year to the next. Tentative evidence is provided suggest-

TABLE 1.12: CIT AND REPORTING GAPS: TRADE IN SERVICES (BOTH SIDES)

	Dep. variable: Value gap		
	OECD sample		
	(1) All flows	(2) 1-digit	(3) 2-digit
CIT Exporter	-0.0141***	-0.0117**	-0.0161***
	(-3.16)	(-2.53)	(-2.79)
CIT Importer	0.00627	0.00750*	0.00500
	(1.50)	(1.68)	(0.87)
Adjusted R^2	0.582	0.529	0.610
Within R^2	0.004	0.004	0.005
Observations	95,654	30,342	29,526
	within-EU sample		
CIT Exporter	-0.0188***	-0.0171***	-0.0203***
	(-3.95)	(-3.33)	(-3.54)
CIT Importer	0.00839*	0.00979*	0.00792
	(1.82)	(1.96)	(1.37)
Adjusted R^2	0.608	0.567	0.632
Within R^2	0.004	0.004	0.005
Observations	102,612	29,634	33,023
Year FE	✓	✓	✓
Importer-exporter-product FE	✓	✓	✓
Controls	✓	✓	✓

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: OECD and EU countries. CIT refers to the central CIT. The columns refer to different levels of aggregation: (1) all flows; (2) 1-digit flows (12 categories); (3) 2-digit flows (32 categories). The within-EU sample is larger due to countries in the EU yet not members of the OECD.

ing that this pattern may be related to tax evasion motives. Second, the tax changes are studied in an event study framework, which allows to check how the reporting gaps behave around the time at which tax rates change.

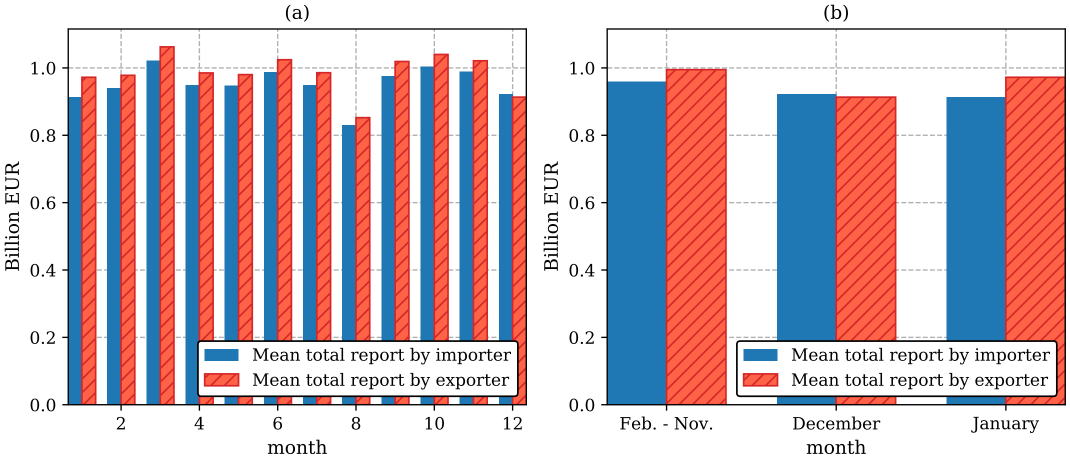
1.5.5.1 Shifting across years

A stark property of the reporting gaps observed at a monthly frequency is that the mean gap is much lower in December than in the other months, and much higher in January (cf. figure 1.B.10).[58] This pattern means that in December, the difference between reported exports and imports is smaller relative to other months, and the other way around in January. Interestingly, the mean gap over the months of December and January *together* is close to that in the other months. It is not obvious that the business cycle or holiday periods can explain this pattern as it is unclear why a firm would re-

---

[58] This is also very noticeable in the time series of the mean gap in figure 1.B.11.

FIGURE 1.1: DECEMBER-JANUARY GAP DIFFERENCE: SEPARATE REPORTS



Note: Panel (a) displays the mean of reports by importers and exporters of total trade between them in a given period (month in a year), by month. Panel (b) displays the same, but the months from February to November are bunched together.

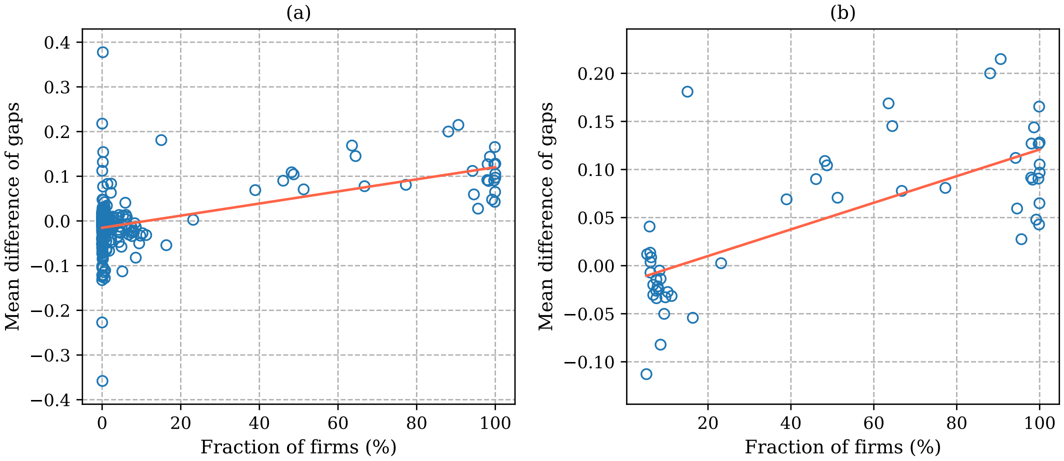
port its imports, but not its exports in December, and conversely in January. A possible explanation is that firms postpone the reporting of their sales to the next year or bring their future imports forward — although it seems more difficult — or both, effectively shifting profits to the next year.[59] Is this pattern related to tax optimization? If the tax schedule is progressive for instance, there is a possible benefit from lowering profits in the current year — if the firm can be taxed at a lower bracket by doing so. Even in the case of a flat tax rate, paying taxes in the future is preferable from a net present value perspective or for liquidity reasons. If that is so, this pattern (a small, perhaps negative gap followed by a large gap) should be more pronounced: (i) in the months in which a significant fraction of the firms have their financial year-end; (ii) in countries that have a progressive tax schedule; (iii) in the months immediately surrounding a tax decrease — and conversely in case of a tax increase.

First, it would be ideal to understand whether exports reports are delayed, or imports reports are brought forward. The latter seems less likely, as future imports are not yet realized. While it is difficult to disentangle these effects, an attempt is made by comparing the value of the reports by importers and exporters (of the same trade flow) in December and January, relative to each other and to the other months — that are less or not subject to shifting. The results are shown in figure 1.1. The means of

---

[59] *A priori*, one could also imagine that firms may want to bring sales from the next year forward to the current year and postpone their costs to the next year to increase their profits and thus the performance of the firm in the current year. This could be especially true for listed firms. This is however inconsistent with the observed pattern. Furthermore, there seems to be no relationship between the December-January difference in gaps and the importance of listed firms in a given sector, as depicted in figure 1.B.13.

FIGURE 1.2: TIME DIFFERENCE IN GAPS VERSUS YEAR-END MONTH OF ACCOUNTS



Note: Scatter plots of mean difference in the reporting gaps between two consecutive months (e.g. between December and January) in a country, against the fraction of firms that have their financial year-end in the earlier month (e.g. December). Each dot represents a country-month pair. Gaps and fractions of firms are matched on the exporting country, but the results are unchanged if the importing country is used. Panel (a) does not impose any restrictions on the data, whereas panel (b) only shows data where the fraction of firms is higher than 5%. The account closing date is taken from BVD Amadeus in year 2015.

total trade at the country-pair level in a given period are displayed for each month, as reported by the importer and the exporter. From panel (a), it seems that trade volumes are lower in December and January than in the other months — except August, when trade significantly slows down. The levels of reported imports are almost identical in December and January, yet reported exports are higher in January than they are in December. This would suggest that exports are shifted from one year to the next.

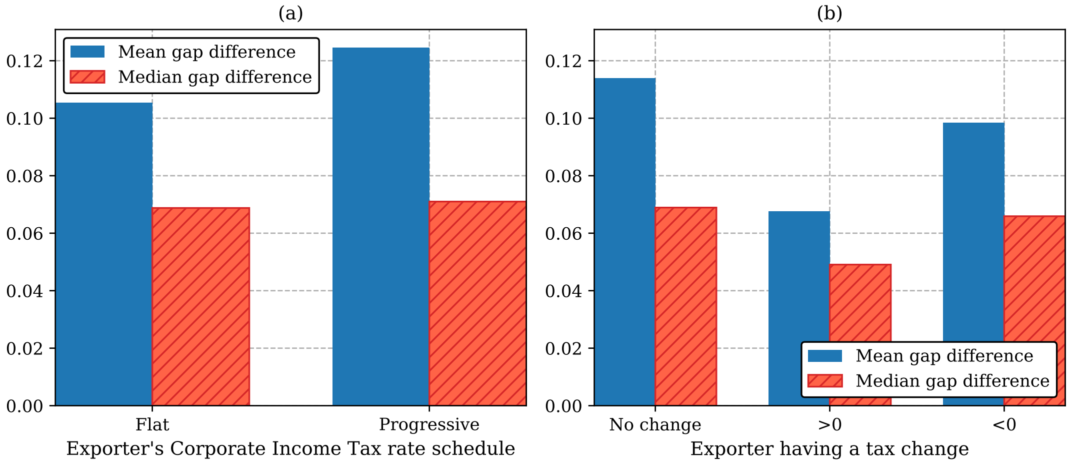
Suggestive evidence for conjecture (i) is provided using data from Amadeus. The fraction of firms having their financial year-end in each months in each country is calculated and plotted against the mean first difference of the reporting gaps in the subsequent month. For instance, the fraction of firms ending their financial year in December in Germany is plotted against the mean difference in gaps between December and January the next year in Germany. The result can be found in figure 1.2. The relation between the two is positive: when more firms have their financial year-end in a given month, the difference in gaps between that month and the following is higher, consistent with the hypothesis that firms delay profits at that crucial time.[60]

If point (ii) is true, one should see a higher difference between the December and January gaps in exporting countries where the tax schedule is progressive. It is how-

---

[60] The exception seems to be the UK in December-January: although few firms end their financial year then — probably because the fiscal year runs from April to March — the difference in gaps is very large. It is however even larger between March and April, which is not the case in other countries.

FIGURE 1.3: DECEMBER-JANUARY GAP DIFFERENCE AND TAX RATES



Note: Panel (a): Mean and median December-January difference in gaps, for exporters that have a flat and a progressive tax schedule. The countries with a progressive tax as of 2018 are listed in table 1.A.2. Reporting gaps are net of product and importer-year fixed effects (exporter fixed effects cannot be included as they would absorb the differences in tax schedules, which are exporter-specific). Panel (b): Mean and median December-January difference in gaps in exporting countries where: (i) there was no tax change, (ii) there was a positive tax change, (iii) there was a negative tax change. Reporting gaps are net of importer-exporter-product and importer-year fixed effects.

ever not obvious this effect will be detectable, as the firms responsible for most of international trade are large firms, whose taxable profits are more likely to be large and therefore significantly past the threshold of the highest tax bracket. Comparing the mean differences in gaps of exporting countries that have a progressive tax schedule versus not suggests that there is less shifting in flat-tax countries relative to countries that have a progressive rate, in line with the hypothesis — see panel (a) of figure 1.3.[61]

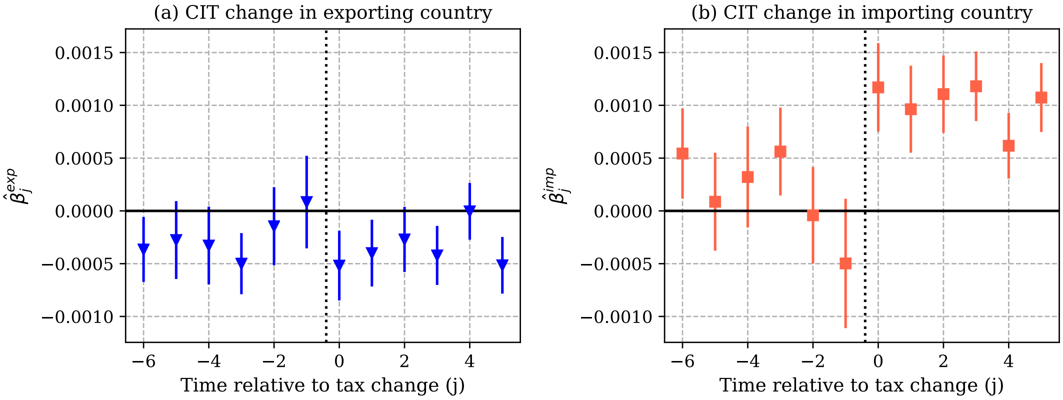
Conjecture (iii) stipulates that if a firm knows that it will face a higher tax rate the following year, the December-January difference in the gaps should be less pronounced, as less profits should be shifted from the current year to the next. In order to tentatively support this claim, the mean and median difference in gaps between December and January is taken for exporting countries and years when there is no tax change and is compared to episodes of tax increase and decrease. The results are displayed panel (b) of figure 1.3.[62] As expected, the differences in gaps are smaller when there is a positive tax change, consistent with firms wishing to shift exports less when the tax rate increases (i.e. firms shift profits to a lesser extent when the tax rate increases in the following year). However, the gap difference is also slightly lower when a negative tax change takes place, which is inconsistent with the conjecture. It is

---

[61] Countries' tax regime (flat versus progressive) is as of 2018. This exercise does not take into account countries that switched regimes over time and should therefore be taken with the necessary prudence.

[62] In panel (b) of this figure, the gaps are net of importer-exporter-product and importer-year fixed effects. The patterns are similar if the raw gaps are used.

FIGURE 1.4: MONTHLY REPORTING GAPS AND CIT CHANGE



Note: Estimates of β_j from regression equation (1.18), for tax change in the importing country (i.e. $c = i$, depicted in orange) and in the exporting country (i.e. $c = e$, depicted in blue). The sample is the within-EU sample. Data at a monthly frequency.

however reassuring to see that when exporters face a tax increase, the shift in profits seems *significantly* smaller — the difference in gaps is more than 40% smaller in these instances.

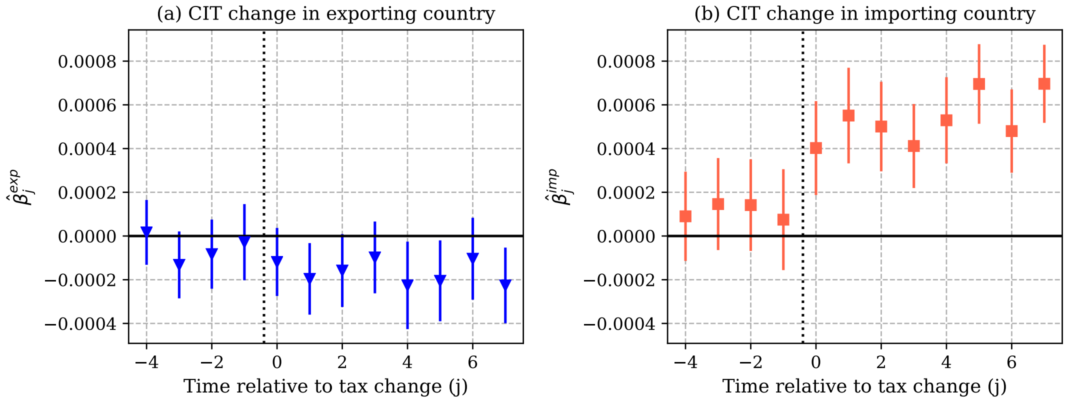
In summary, it appears that some exports that accrue to December may be shifted to the next calendar year. It is not possible to say with certainty that this shifting is related to tax purposes, but it may. First, this shifting seems more prevalent in months when more firms have their financial year-end. Second, it appears to be less prevalent when taxes increase in the following year, and when the tax rate schedule is flat.

1.5.5.2 Event study

The event study intends to refine the identification of corporate tax evasion by looking at a narrowly defined window around precisely dated tax changes using data at a monthly frequency. This arguably reduces the likelihood of the results being driven by an omitted variable. However, virtually all tax changes take place on January 1st, which is not an innocuous date: many legal and administrative changes take place at the start of the year. The estimated coefficients may be capturing these changes if they affect the reporting gaps. Furthermore, it is not clear that firms will immediately adjust their reported sales and purchases already in January since profits of the whole year are taxed once — although it is possible that firms adjust gradually to avoid detection by authorities via their VAT returns.

Estimates of regression (1.18) are provided, using monthly data and quarterly data. Quarterly data is less noisy and the estimates are more stable. It also allows

FIGURE 1.5: QUARTERLY REPORTING GAPS AND CIT CHANGE


(a) CIT change in exporting country

(b) CIT change in importing country

Note: Estimates of β_j from regression equation (1.18), for tax change in the importing country (i.e. $c = i$, depicted in orange) and in the exporting country (i.e. $c = e$, depicted in blue). The sample is the within-EU sample. Data at the quarterly frequency.

the study the dynamic effects over longer periods, which is computationally difficult using monthly data — including lags and leads in equation (1.18) is very expensive. All regressions control for the standard VAT rate and the Intrastat reporting thresholds (cf. section 1.3.2.2) as well as real GDP. The model and previous results using annual data suggest that the estimated β_j^{exp} should be negative after the tax change — as exporters lower their reported exports when the tax rate increases. The series of β_j^{imp} could either be positive or negative, but previous analysis on within-EU trade using yearly data suggests they should be positive. Figures 1.4 and 1.5 depict the coefficients β_j around tax changes in the importing and exporting countries at the monthly and quarterly frequencies, respectively. In figure 1.4, the horizon is 6 months at either side of the tax change, whereas in figure 1.5, the window spans 1 year (4 quarters) before the change, and 2 years (8 quarters) after the change.

Some patterns emerge from the figures. First, at the monthly frequency not all coefficients are insignificant before the event date, which may be indicative of differential pre-trends. This however disappears when using quarterly data. Second, the dynamic effects following tax changes have the expected sign in all specifications. Third, the estimates $\hat{\beta}_j^{exp}$ are smaller in magnitude than $\hat{\beta}_j^{imp}$, which suggests that firms under-report costs *more* that they do sales, which is at odds with previous results and counter-intuitive as it could result in an increase in taxable profits (only if purchases are large enough relative to sales). Last, it appears that the coefficients on CIT are much larger than in the yearly regressions. In the yearly sample, the estimated semi-elasticities of the reporting gaps with respect to the CIT range from 0.27% to 0.29% in magnitude (see table 1.7), whereas they range from 0.02% to 0.12% in the event study. This is a large difference which may stem from differences in the data on gaps or perhaps because firms do not adjust completely their behaviour within the time window considered in the event study. Overall, this exercise is interesting for two reasons: first, it suggests that the parallel pre-trend assumption seems to hold. Second, the gaps react in the expected way shortly after the tax rate change.

1.6 Tariff evasion

This section details evidence of tariff evasion. First, I first present the baseline result and put it in relation to existing estimates from the literature. Second, tariff evasion is shown to be more important for differentiated products, in high-corruption and lower-income countries, and appears to be constant over time. Evasion also occurs through product misclassification. These results largely echo previous findings and add new insights, leveraging the extensive data coverage.

Baseline result Table 1.A.18 reports the effects of tariffs on the reporting gaps, as defined earlier in the paper. The dependent variables are the reporting gaps in log in terms of value and quantities, and the coefficient of interest is that on the tariff the importer faces, α in equation (1.14). Every regression includes the most restrictive set of fixed effects (FE): importer-exporter-product and importer-exporter-year. Starting from column (1), a 1 percentage point increase in the tariff rate is associated with an increase in the value gap by 0.16%. This is consistent with the model prediction that importers under-report imports as the tariff rate increases — cf. prediction 1. Column (2) suggests that evasion is mostly taking place through manipulation of prices as opposed to quantities. The within R-squared is close to zero, indicating that changes in tariffs do not explain much of the variation in the gaps overall, consistent with the prior that the reporting gaps are subject to considerable measurement errors. Decomposing the effect further, it appears that as the tariff increases, both reported imports and exports decrease — as one would expect, a higher tariff has a real negative effect on trade — yet reported imports decreases by more than its counterpart reported by the exporter, see columns (1) and (2) of table 1.A.22. The effect of tariffs on the reporting gaps is almost linear: the coefficient on tariff squared is negative and significant, yet tiny in magnitude — see table 1.A.23. The baseline semi-elasticity of 0.16% is in line with — yet below what — other studies using the same method have generally found — see table 1.A.24. It is however difficult to compare results across studies, as each focused on distinct countries in specific episodes. The finding that tariff evasion does not appear to take place through quantities manipulation is puzzling, since most countries in the sample are part of the WTO and are thus bound to use the same valuation method to assess the dutiable value of imports, which has been shown to make price manipulation more difficult (Javorcik and Narciso, 2017).[63] In this sample, evasion levels are almost identical whether the importing country is in the WTO or not, as reported in table 1.A.25.

Including such a battery of fixed effects in the regressions may be unnecessarily absorbing much of the variation in the data. Whether it is indeed the correct strategy cannot be said for certain, as dropping fixed effects leaves room for potential omitted variable biases. However, the distributions of the reporting gaps suggest that OVB might be a concern and warrants the use of fixed effects. Indirect evidence that these fixed effects are relevant is provided in table 1.A.26. The coefficient on tariff decreases markedly as more fixed effects are added progressively. The largest decrease in the estimated α takes place as importer-exporter-product fixed effects are included.[64]

Differentiated products A differentiated product is defined as a good for which no organized market nor reference price exist. They are in principle harder to value, which may leave room for traders to misreport the unit value of the good they trade (Fisman and Wei, 2009; Javorcik and Narciso, 2008). Tariff evasion is typically found to be higher for differentiated products in the literature (Javorcik and Narciso, 2008). In this sample, this finding only holds true when a subset of the fixed effects are included. The first two columns of table 1.A.19 refer to regressions where the importer-exporter-product fixed effect is replaced by product a fixed effect only. In those specifications: (i) evasion is higher for differentiated products; (ii) evasion takes place through quantities mis-reporting when products are not differentiated, and mostly through prices when they are. Those results echo those of Javorcik and Narciso (2008), with the difference that evasion through quantities mis-reporting seems to exist for

---

[63] The implementation of the WTO Customs Valuation Agreement (CVA) has been shown to displace evasion from price manipulation to quantities manipulation in the case of differentiated products. More details on the CVA are given in appendix 1.D.

[64] In the case of regressions of gaps in terms of quantities, the signs even switch sign in some instances. Generally, the data on quantities seems of lesser quality, as suggested in section 1.3. The regression coefficients are also generally less precisely estimated and less stable across specifications.

non-differentiated products in this sample.[65] However, these effects disappear once (i, e, p) fixed effects are included in the regression.

Corruption The model predicts that lower efficiency of the tax authorities should result in more evasion, and therefore a stronger reaction of the reporting gaps to changes in tax rates — cf. prediction 4. In the case of tariffs, customs are the relevant authorities, and lower quality customs should result in higher evasion. Two measures of customs efficiency are provided by the World Bank, but they only exist since 2007. A corruption index by the International Country Risk Guide (ICRG), is used as a proxy.[66] Equation (1.14) is estimated, including an interaction of tariff and a dummy indicating whether the importing country has high corruption levels.[67] Alternatively, a dummy indicating that both the importing and exporting countries have high levels of corruption is used in the regressions. The results can be found in table 1.A.20. Tariff evasion is higher when the importing country has high corruption, and also when both trading partners are countries with high corruption levels. These results are consistent with model prediction 4, if one interprets corruption levels as a proxy for model parameter b in the cost function.

Misclassification Another type of misreporting studied in the literature is misclassification across product categories (see, e.g. Fisman and Wei, 2004; Javorcik and Narciso, 2008). The intuition is as follows: if the tariffs levied on a similar product decreases, a trader may be tempted to misclassify its trade as that product with a lower tariff. Reported imports of the *correct* product decrease and the reporting gap of that product increases. This suggests a negative relationship between tariffs on similar products and the reporting gaps. Regression (1.14) is augmented with the average tariff of other products within the same 4- or 5-digit product group.[68] Table 1.A.21

---

[65] Javorcik and Narciso (2008) do not find evidence of evasion through quantities mis-reporting in their sample, which covers trade between Germany and 10 Eastern European countries during 1992–2003. The liberal classification of Rauch (1999) is used in this study, but the results remain unchanged when using the conservative version of the classification.

[66] The two measures are: *Efficiency of customs clearance process* (series LP.LPI.CUST.XQ) and *Burden of customs procedure, WEF* (series IQ.WEF.CUST.XQ). These measures and the corruption index are highly correlated (correlation coefficients of 0.8 and 0.75 across countries over 2007-2017, respectively).

[67] The corruption index ranges from 0 to 6, where 0 is the highest level of corruption. The high corruption dummy takes on value 1 if the index is below 3, which is the median value of the index in the world sample.

[68] The average tariff is calculated as a weighted average of the tariffs of all other products within the same 4- or 5-digit product group, using the value of trade — as measured by the reported exports — as weights. Formally: $\bar{t}_{-p} = \sum_{k \in P \setminus \{p\}} t_k \theta_k$, where P denotes the set of products within the product group and

displays the results. The coefficients on tariffs on similar products are negative, consistent with product misclassification. This phenomenon appears to be happening through quantities, consistent with the intuitive idea that whole transactions are misclassified.[69]

Evasion over time and across income levels By interacting tariffs with indicators of 5-year time periods as well as income levels, it is possible to investigate whether tariff evasion has changed over time, and whether it is more prevalent in less-developed countries. Figure 1.B.16 displays these estimates.[70] The reaction of reporting gaps to tariffs is very stable over time apart for the period 1988-1993, possibly due to the poor quality and coverage of the sample in these early years. Tariff evasion seems more prevalent in low income countries. The coefficient on tariffs is even negative in high income countries, which does not align with model predictions. It is puzzling and somewhat worrisome that the coefficient is *significantly* negative.

Estimated amounts of evasion The estimated reaction of reporting gaps to changes in tariffs is interpreted as evidence of evasion. Based on the estimates above, a tentative figure of the amounts evaded for a country i in year t can be calculated as follows:

$$\text{Amount evaded}_{it} = \sum_{p} \text{Trade flow}_{ipt} \times \text{tariff}_{ipt} \times \hat{\alpha}_G, \tag{1.20}$$

where $\hat{\alpha}_G$ is the estimated coefficient on tariff for importer in income group $G \in \{L, LM, UM\}$, i.e. low (L), lower-middle (LM) and upper-middle (UM) income groups. The negative estimated α in high income countries is interpreted as evidence that no tariff evasion takes place in these countries. Expression (1.20) implicitly assumes that the marginal evasion is constant over time — which is supported by figure 1.B.16, panel (a) — and for any value of the tariff within income group G. Another more subtle assumption implicitly made is that imports are under-reported and not misclassified as another product upon an increase in tariffs. If misclassification occurs — as evidenced above — evasion amounts as calculated in (1.20) will be an

---

$\theta_k = \text{exports}_k / \sum_{j \in P \backslash \{p\}} \text{exports}_j$ is the ratio of the value of trade of product k over total trade of products within the group other than product p, as measured by the reported exports. This is the measure used in other studies (Fisman and Wei, 2004; Javorcik and Narciso, 2008).

[69] Misclassification through prices would imply manipulating prices across trade flows of similar products, which is less intuitive and could only happen when a transaction is composed of at least two similar goods, and the importer manipulates the prices on the invoice.

[70] The underlying regression results can be found in tables 1.A.27 and 1.A.28.

upper bound for the true levels of evasion and should therefore be treated as such. Imports are used to measure the trade flow.[71] Globally, evasion amounts (at most) to an average of 44 Billion US dollars (USD) annually over the last 10 years, as depicted in figure 1.B.17. This represents about 2.5% of the flows subject to tariffs. To give context to the estimated amount of evasion, it can be expressed as percentage of actual import duties revenues.[72] This data is not consistently available for every country over the sample period. To avoid composition changes over time, this measure is only computed for 19 countries over the years 2002-2016 for which import duties revenues are consistently available.[73] These are mostly less-developed countries where tariff levels are generally higher and tariff revenues more important relative to higher-income countries. On average, tariff evasion represents around 32% of actual revenues in these countries, a substantial amount.[74]

1.7 Conclusion

I provide novel evidence of corporate income tax evasion employing a method originally used to estimate tariff evasion. Evasion strategies of the firm are motivated in an illustrative model that delivers predictions which are intuitive, yet non-trivial. Upon an increase in the corporate income tax rate, firms under-report exports to lower their taxable profits and may wish to under- or over-report imports depending on institutional features, the level of tariffs and the extent to which costs are tax-deductible. Empirical identification of evasion relies on how the discrepancies between reports by exporters and importers of the same trade flow react to changes in the statutory corporate income tax rates faced by importers and exporters. The data supports each model prediction, lending strength to a causal interpretation of the results. Evasion is lower when firms are subject to stricter anti-evasion policies. Monthly and quar-

---

[71] The drawback of using imports is that imports and tariffs are related through evasion. Since higher tariffs imply evasion and thus lower imports, using imports to measure the trade flow will underestimate the amounts evaded. An alternative is to use the average of the components of the reporting gaps. This is however only possible for trade flows where both sides of the gap are available, which only represents about 77% of world trade.

[72] The data on *Customs and other import duties* (ID: GC.TAX.IMPT.CN) comes from the World Bank.

[73] The countries in the sub-sample are: Argentina, Bangladesh, Armenia, Brazil, Costa Rica, El Salvador, Georgia, Guatemala, Jordan, Korea, Malaysia, Mauritius, Moldova, Morocco, Namibia, Nicaragua, Peru, South Africa and Uruguay.

[74] This figure is 37% based on all observations for which data is available — dropping Bosnia and Herzegovina, Slovak Republic, Estonia and Cape Verde for which the data on imports duties revenues is very poor (very large jumps and negative values). Note however that the data on import duties revenues is not available for every country and for every year. In particular, it is unavailable for many high-income countries.

terly data is used to provide further evidence of evasion by studying the response of trade discrepancies to tax changes around the date at which rates change in an event study framework. Patterns suggesting that firms shift profits across years are unveiled, along with evidence that these may be related to tax evasion motives.

1.8 Bibliography

Abraham, Sarah and Liyang Sun, "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," 2018.

Allingham, Michael G and Agnar Sandmo, "Income tax evasion: A theoretical analysis," *Journal of public economics*, 1972, *1* (3-4), 323–338.

Alm, James, Jorge Martinez-Vazquez, and Chandler McClellan, "Corruption and firm tax evasion," *Journal of Economic Behavior & Organization*, 2016, *124*, 146–163.

Almunia, Miguel and David Lopez-Rodriguez, "Under the radar: The effects of monitoring firms on tax compliance," *American Economic Journal: Economic Policy*, 2018, *10* (1), 1–38.

Becker, Johannes and Clemens Fuest, "Optimal tax policy when firms are internationally mobile," *International Tax and Public Finance*, 2011, *18* (5), 580–604.

Best, Michael Carlos, Anne Brockmeyer, Henrik Jacobsen Kleven, Johannes Spinnewijn, and Mazhar Waseem, "Production versus revenue efficiency with limited tax capacity: theory and evidence from Pakistan," *Journal of political Economy*, 2015, *123* (6), 1311–1355.

Bilicka, Katarzyna Anna, "Comparing UK tax returns of foreign multinationals to matched domestic firms," *American Economic Review*, 2019.

Borusyak, Kirill and Xavier Jaravel, "Revisiting event study designs," *Available at SSRN 2826228*, 2017.

Braml, Martin T and Gabriel J Felbermayr, "What Do We Really Know about the Transatlantic Current Account?," *CESifo Economic Studies*, 09 2019, *65* (3), 255–274.

Bussy, Adrien, "Cross-border Value Added Tax Fraud In The European Union," Working Paper 2020.

Carrillo, Paul, Dina Pomeranz, and Monica Singhal, "Dodging the taxman: Firm misreporting and limits to tax enforcement," *American Economic Journal: Applied Economics*, 2017, *9* (2), 144–64.

Chetty, Raj, John N Friedman, Tore Olsen, and Luigi Pistaferri, "Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from Danish tax records," *The Quarterly Journal of Economics*, 2011, *126* (2), 749–804.

Clausing, Kimberly A, "Corporate tax revenues in OECD countries," *International tax and public finance*, 2007, *14* (2), 115–133.

DeBacker, Jason, Bradley T Heim, and Anh Tran, "Importing corruption culture from overseas: Evidence from corporate tax evasion in the United States," *Journal of Financial Economics*, 2015, *117* (1), 122–138.

Demir, Banu and Beata Javorcik, "Trade policy changes, tax evasion and Benford's law," *Journal of Development Economics*, 2020, p. 102456.

Devereux, Michael P and Rachel Griffith, "Taxes and the Location of Production: Evidence from a Panel of US Multinationals," *Journal of public Economics*, 1998, *68* (3), 335–367.

_ **and** _ , "Evaluating tax policy for location decisions," *International tax and public finance*, 2003, *10* (2), 107–126.

_ **, Li Liu, and Simon Loretz**, "The elasticity of corporate taxable income: New evidence from UK tax records," *American Economic Journal: Economic Policy*, 2014, *6* (2), 19–53.

_ **, Rachel Griffith, and Alexander Klemm**, "Corporate income tax reforms and international tax competition," *Economic policy*, 2002, *17* (35), 449–495.

Dowd, Tim, Paul Landefeld, and Anne Moore, "Profit shifting of US multinationals," *Journal of Public Economics*, 2017, *148*, 1–13.

Drechsler, Itamar, Alexi Savov, and Philipp Schnabl, "The deposits channel of monetary policy," *The Quarterly Journal of Economics*, 2017, *132* (4), 1819–1876.

Dutt, Pushan and Daniel Traca, "Corruption and bilateral trade flows: extortion or evasion?," *The Review of Economics and Statistics*, 2010, *92* (4), 843–860.

Dwenger, Nadja and Viktor Steiner, "Profit Taxation and the Elasticity of the Corporate Income Tax Base: Evidence From German Corporate Tax Return Data," *National Tax Journal*, 2012, *65* (1), 118–50.

_ **and** _ , "Financial leverage and corporate taxation: Evidence from German corporate tax return data," *International tax and public finance*, 2014, *21* (1), 1–28.

European Commission, "The concept of Tax Gaps. Report II: Corporate Income Tax Gap Estimation Methodologies," July 2018.

_ , "The concept of Tax Gaps. Report III: MTIC Fraud Gap estimation methodologies," November 2018.

Eurostat, "User guide on European statistics on international trade in goods. 2016 Edition," 2016.

_ , "National requirements for the Intrastat system. 2018 Edition," 2017.

Farhad, Mohammad, Michael Jetter, Abu Siddique, and Andrew Williams, "Misreported Trade," 2018.

Ferrantino, Michael J, Xuepeng Liu, and Zhi Wang, "Evasion behaviors of exporters and importers: Evidence from the US–China trade data discrepancy," *Journal of International Economics*, 2012, *86* (1), 141–157.

Fisman, Raymond and Shang-Jin Wei, "Tax rates and tax evasion: evidence from "missing imports" in China," *Journal of political Economy*, 2004, *112* (2), 471–496.

_ **and** _ , "The smuggling of art, and the art of smuggling: Uncovering the illicit trade in cultural property and antiques," *American Economic Journal: Applied Economics*, 2009, *1* (3), 82–96.

Francois, Joseph and Olga Pindyuk, "Consolidated Data on International Trade in Services v8. 9," *IIDE Discussion Paper 20130101*, 2013.

Frunza, Marius, "Cost of the MTIC VAT fraud for European Union members," *Available at SSRN 2758566*, 2016.

Fuest, Clemens, Andreas Peichl, and Sebastian Siegloch, "Do higher corporate taxes reduce wages? Micro evidence from Germany," *American Economic Review*, 2018, *108* (2), 393–418.

Gemmell, Norman and John Hasseldine, "Taxpayers' behavioural responses and measures of tax compliance 'gaps': A critique and a new measure," *Fiscal Studies*, 2014, *35* (3), 275–296.

Gradeva, Katerina, "VAT Fraud in Intra-EU Trade," *Goethe University, Frankfurt*, 2014.

Gruber, Jonathan and Joshua Rauh, "How elastic is the corporate income tax base?," *Taxing corporate income in the 21st century*, 2007, pp. 140–163.

Hebous, Shafik and Niels Johannesen, "At your service! The role of tax havens in international trade with services," 2015.

IRS, "Federal Tax Compliance Research: Tax Gap Estimates for Tax Years 2008–2010," Technical Report, Internal Revenue Service Research, Analysis and Statistics 2016.

Javorcik, Beata S and Gaia Narciso, "Differentiated products and evasion of import tariffs," *Journal of International Economics*, 2008, *76* (2), 208–222.

_ and _ , "WTO accession and tariff evasion," *Journal of Development Economics*, 2017, *125*, 59–69.

Joulfaian, David, "Corporate income tax evasion and managerial preferences," *Review of Economics and Statistics*, 2000, *82* (4), 698–701.

Kawano, Laura and Joel Slemrod, "How do corporate tax bases change when corporate tax rates change? With implications for the tax rate elasticity of corporate tax revenues," *International Tax and Public Finance*, 2016, *23* (3), 401–433.

Kee, Hiau Looi and Alessandro Nicita, "Trade frauds, trade elasticities and non-tariff measures," in "5th IMF-World Bank-WTO trade research workshop," Vol. 30 2016.

Keen, Michael and Joel Slemrod, "Optimal tax administration," *Journal of Public Economics*, 2017, *152*, 133–142.

Kellenberg, Derek and Arik Levinson, "Misreporting trade: Tariff evasion, corruption, and auditing standards," *Review of International Economics*, 2019, *27* (1), 106–129.

Kleven, Henrik J and Mazhar Waseem, "Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan," *The Quarterly Journal of Economics*, 2013, *128* (2), 669–723.

Kleven, Henrik Jacobsen, Martin B Knudsen, Claus Thustrup Kreiner, Søren Pedersen, and Emmanuel Saez, "Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark," *Econometrica*, 2011, *79* (3), 651–692.

Levin, Jörgen and Lars M Widell, "Tax evasion in Kenya and Tanzania: Evidence from missing imports," *Economic Modelling*, 2014, *39*, 151–162.

Liu, Xuepeng, "Tax avoidance through re-imports: The case of redundant trade," *Journal of Development Economics*, 2013, *104*, 152–164.

Markhonko, Vladimir, "Asymmetries in official international trade statistics and analysis of globalization," *Aguascalientes, Mexico*, 2014.

Miao, Guannan and Fabienne Fortanier, "Estimating Transport and Insurance Costs of International Trade," 2017.

Mishra, Prachi, Arvind Subramanian, and Petia Topalova, "Tariffs, enforcement, and customs evasion: Evidence from India," *Journal of public Economics*, 2008, *92* (10), 1907–1925.

Naritomi, Joana, "Consumers as tax auditors," *American Economic Review*, 2019, *109* (9), 3031–72.

OECD, "OECD tax database: Explanatory annex," Technical Report 2019.

Rauch, James E, "Networks versus markets in international trade," *Journal of international Economics*, 1999, *48* (1), 7–35.

Saez, Emmanuel, "Do taxpayers bunch at kink points?," *American economic Journal: economic policy*, 2010, *2* (3), 180–212.

_ , Joel Slemrod, and Seth H Giertz, "The elasticity of taxable income with respect to marginal tax rates: A critical review," *Journal of economic literature*, 2012, *50* (1), 3–50.

Schmidheiny, Kurt and Sebastian Siegloch, "On Event Study Designs and Distributed-Lag Models: Equivalence, Generalization and Practical Implications," 2019.

Sequeira, Sandra, "Corruption, trade costs, and gains from tariff liberalization: evidence from Southern Africa," *The American Economic Review*, 2016, *106* (10), 3029–3063.

Serrato, Juan Carlos Suárez and Owen Zidar, "Who benefits from state corporate tax cuts? A local labor markets approach with heterogeneous firms," *American Economic Review*, 2016, *106* (9), 2582–2624.

Serrato, Juan Carlos Suárez and Owen Zidar, "The structure of state corporate taxation and its impact on state tax revenues and economic activity," *Journal of Public Economics*, 2018, *167*, 158–176.

Slemrod, Joel, "Tax compliance and enforcement: New research and its policy implications," 2016.

_ , Brett Collins, Jeffrey L Hoopes, Daniel Reck, and Michael Sebastiani, "Does credit-card information reporting improve small-business tax compliance?," *Journal of Public Economics*, 2017, *149*, 1–19.

Spengel, C, F Schmidt, JH Heckemeyer, K Nicolay, A Bartholmeß, C Ludwig, D Steinbrenner, R Bräutigam, P Buchmann, T Bührle et al., "Effective Tax Levels Using the Devereux/Griffith Methodology-Update 2018," *Project for the EU Commission TAXUD/2018/DE/307 Intermediary Report 2018, Mannheim*, 2019.

Stoyanov, Andrey, "Tariff evasion and rules of origin violations under the Canada-US Free Trade Agreement," *Canadian Journal of Economics/Revue canadienne d'économique*, 2012, *45* (3), 879–902.

Sullivan, Clare Linda and Evan Smith, "Trade-based money laundering: Risks and regulatory responses," *Australian Institute of Criminology*, 2012, pp. 1–27.

Tørsløv, Thomas R, Ludvig S Wier, and Gabriel Zucman, "The missing profits of nations," Technical Report, National Bureau of Economic Research 2018.

Ueda, Junji, "Estimating the Corporate Income Tax Gap: The RA-GAP Methodology," Technical Report, IMF August 2018.

UN, "The International Merchandise Trade Statistics: Concepts and Definitions 2010," Technical Report, Department of Economic and Social Affairs, United Nations 2011.

_ , "IMTS Bilateral asymmetries – how to measure, analyze, reduce and way forward," Technical Report, Department of Economic and Social Affairs, United Nations 2019.

UNSD, *A Manual on Statistics of International Trade in Services 2002*, Vol. 86, International Monetary Fund, 2002.

_ , "Data processing in UN Comtrade," 2017. United nations statistics division. *UNSD*

UNSD, *https://unstats.un.org/unsd/tradekb/knowledgebase/ 2019. Accessed: 16-07-2019.*

UNSD, *"IMTS Bilateral asymmetries – how to measure, analyze, reduce and way forward," 2019. United nations statistics division.*

Yang, Dean, *"Can enforcement backfire? Crime displacement in the context of customs reform in the Philippines," The Review of Economics and Statistics, 2008, 90 (1), 1–14.*

Yitzhaki, Shlomo, *"A note on income tax evasion: A theoretical analysis.," Journal of public economics, 1974, 3, 201–202.*

Appendix 1.A Additional tables

TABLE 1.A.1: PROVIDERS OF STATISTICAL INFORMATION AND VAT-REGISTERED TRADERS IN 2017

	Intra-EU Imports			Intra-EU Exports		
	PSIs	VAT-reg.	PSI as % of VAR-reg.	PSIs	VAT-reg.	PSI as % of VAR-reg.
Belgium	8,360	267,839	3.1	8,322	63,110	13.2
Bulgaria	7,026	37,692	18.6	5,883	16,916	34.8
Czech Republic	11,134	104,009	10.7	8,879	54,861	16.2
Denmark	5,745	43,369	13.2	4,501	16,430	27.4
Germany	42,334	547,615	7.7	42,605	224,995	18.9
Estonia	3,883	19,173	20.3	3,192	10,616	30.1
Ireland	5,145	48,310	10.6	2,249	12,037	18.7
Greece	7,290	67,141	10.9	4,431	24,622	18.0
Spain	17,700	318,290	5.6	21,500	146,098	14.7
France	42,552	434,416	9.8	25,190	144,289	17.5
Croatia	5,372	199,666	2.7	3,325	65,218	5.1
Italy	78,000	353,000	22.1	70,000	161,000	43.5
Cyprus	2,151	12,457	17.3	357	2,869	12.4
Latvia	4,531	25,337	17.9	3,567	10,446	34.1
Lithuania	4,495	24,101	18.7	3,629	12,683	28.6
Luxembourg	3,400	19,000	17.9	1,400	5,500	25.5
Hungary	6,700	78,900	8.5	5,800	38,400	15.1
Malta	2,911	8,892	32.7	255	2,698	9.5
Netherlands	13,000	279,000	4.7	12,000	112,000	10.7
Austria	10,004	148,780	6.7	5,990	36,196	16.5
Poland	11,481	114,511	10.0	12,062	84,851	14.2
Portugal	11,011	113,739	9.7	7,460	30,339	24.6
Romania	14,838	76,514	19.4	5,700	18,705	30.5
Slovenia	6,949	34,761	20.0	3,450	25,906	13.3
Slovakia	10,898	78,509	13.9	4,657	25,269	18.4
Finland	5,200	74,700	7.0	2,380	14,190	16.8
Sweden	7,567	115,070	6.6	5,984	30,523	19.6
United Kingdom	10,288	155,435	6.6	21,018	115,152	18.3
Average			12.6			20.2

Note: Number of Providers of Statistical Information (PSI) and VAT-registered traders by EU country in 2017. Source: table 1 in Eurostat (2017).

TABLE 1.A.2: FISCAL INFORMATION FOR EU COUNTRIES

Country	Fiscal year	Firms can choose different accounting year	Tax rate structure (main tax, as of 2018)
Austria	calendar year	yes	flat
Belgium	calendar year	yes	2 brackets
Bulgaria	calendar year	no	flat
Croatia	calendar year	yes, but must apply	2 rates (one for small firms)
Cyprus	calendar year	unsure, but seems no	flat[a]
Czech Republic	calendar year	yes, must notify	flat
Denmark	calendar year	yes, if justified by special circumstances	flat
Estonia	calendar year (tax due every month)	yes	flat
Finland	calendar year	yes	flat
France	calendar year	yes	several brackets
Germany	calendar year	yes, subject to authorization by the tax authority	flat
Greece	calendar year	yes	flat
Hungary	calendar year	yes	flat[b]
Ireland	calendar year	yes	flat, but industry-dependent
Italy	calendar year	yes	flat[c]
Latvia	calendar year	yes	flat
Lithuania	calendar year	yes, subject to permission	flat, reduced rate for small firms
Luxembourg	calendar year	yes	several brackets
Malta	calendar year	yes, subject to approval	flat
Netherlands	calendar year	seems not	2 brackets
Poland	calendar year	usually calendar	2 brackets
Portugal	calendar year	yes	complex
Romania	calendar year	yes	flat
Slovak Republic	calendar year	yes	flat
Slovenia	calendar year	yes	flat
Spain	calendar year	yes	flat, but different rates apply to certain types of firms.
Sweden	calendar year	yes	flat
United Kingdom	April 1st - March 31st	yes	flat

---

[a] The tax regime is flat, but the change that took place in 2005 was the removal of a 5% surtax that only affected firms with profits in excess of £1 million.

[b] There were two tax rates (a lower one for firms with low profits) until 01/01/2017.

[c] The main corporate tax rate as of 2003 (IRES) is flat. In the earlier years of the sample, other taxes applied, for which a variety of deductions existed.

TABLE 1.A.3: AVERAGE NUMBER OF FIRMS PER YEAR IN AMADEUS SAMPLE

Country	N firms
AT	148,095
BE	419,345
BG	320,010
CY	1,126
CZ	184,544
DE	469,822
DK	254,311
EE	124,572
ES	772,505
FI	172,104
FR	702,746
GB	2,715,271
GR	23,268
HR	105,928
HU	430,139
IE	146,347
IT	938,301
LT	11,745
LU	23,186
LV	112,481
MT	15,275
NL	711,566
PL	165,809
PT	350,281
RO	665,036
SE	441,207
SI	70,720
SK	172,671

Note: The sample is Amadeus, years 2015-2017. Sample restricted to firms for which total assets is non-missing.

TABLE 1.A.4: CORPORATE INCOME TAX RATES AND REPORTING GAPS: EACH SIDE SEPARATELY WITHOUT TIME TRENDS

	(1) Value gap, log	(2) Quantity gap, log	(3) Value gap, log	(4) Quantity gap, log
CIT Exporter	-0.00225***	-0.00156***		
	(-5.13)	(-2.90)		
CIT Importer			0.000626	0.000505
			(1.03)	(0.74)
Tariff	0.000998***	-0.00125***	-0.000804	-0.00242***
	(4.31)	(-2.99)	(-1.53)	(-3.03)
Importer-export-product FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Importer-year FE	✓	✓		
Exporter-year FE			✓	✓
Exporter-specific time trend				
Importer-specific time trend				
Controls	✓	✓	✓	✓
Adjusted R^2	0.325	0.321	0.334	0.328
Within R^2	0.000	0.001	0.000	0.000
Observations	44,623,145	38,666,718	33,866,668	29,732,923

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: OECD countries. CIT refers to the central CIT.

TABLE 1.A.5: CIT AND REPORTING GAPS: AGGREGATED TRADE

	(1) Value gap, log	(2) Value gap, log	(3) Value gap, log	(4) Value gap, log
CIT Exporter	-0.00165***		-0.00147***	
	(-3.52)		(-3.00)	
CIT Importer		0.0000690		0.00206***
		(0.13)		(3.66)
Mean tariff	0.000872***	0.000670***	0.000836***	0.000879***
	(3.66)	(2.76)	(3.53)	(3.75)
Importer-exporter-product FE	✓	✓	✓	✓
Importer-year FE	✓		✓	
Exporter-year FE		✓		✓
Exporter-specific time trend	✓			
Importer-specific time trend		✓		
Controls	✓	✓	✓	✓
Adjusted R^2	0.331	0.327	0.334	0.334
Within R^2	0.000	0.000	0.000	0.000
Observations	3,610,495	3,025,275	3,610,495	3,025,275

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: OECD countries. Trade flows are aggregated to the HS 2-digit level. Columns (1) and (2) include country-specific time trends, whereas columns (3) and (4) do not. Quantities cannot be aggregated, as the measurement unit may differ across goods within an HS2 group.

TABLE 1.A.6: CIT AND REPORTING GAPS: FIRST DIFFERENCE

	(1) D.Value gap, log	(2) D.Quantity gap, log	(3) D.Value gap, log	(4) D.Quantity gap, log
D.CIT Exporter	-0.000998***	0.000147		
	(-3.79)	(0.36)		
D.CIT Importer			0.0000609	-0.00174**
			(0.13)	(-2.40)
D.Tariff	0.000357**	0.000416	0.000546**	0.00468***
	(2.52)	(1.38)	(2.39)	(3.39)
Importer-export-product FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Importer-year FE	✓	✓		
Exporter-year FE			✓	✓
Controls	✓	✓	✓	✓
Adjusted R^2	-0.042	-0.037	-0.039	-0.043
Within R^2	0.000	0.000	0.000	0.000
Observations	31,665,970	26,843,676	25,564,871	21,946,794

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: OECD countries. The operator $D.x$ means that variable x is first-differenced, i.e. $D.x_{j,t} = x_{j,t} - x_{j,t-1}$.

TABLE 1.A.7: CORPORATE INCOME TAX RATES AND REPORTING GAPS: EACH SIDE SEPARATELY, BALANCED SAMPLE

	(1) Value gap, log	(2) Quantity gap, log	(3) Value gap, log	(4) Quantity gap, log
CIT Exporter	-0.00148***	-0.00177***		
	(-3.51)	(-3.13)		
CIT Importer			0.000731	-0.000221
			(1.07)	(-0.30)
Tariff	0.000627**	-0.00183***	-0.000833*	-0.00134*
	(2.43)	(-3.83)	(-1.83)	(-1.75)
Importer-export-product FE	✓	✓	✓	✓
Importer-year FE	✓	✓		
Exporter-year FE			✓	✓
Exporter-specific time trend	✓	✓		
Importer-specific time trend			✓	✓
Controls	✓	✓	✓	✓
Adjusted R^2	0.346	0.334	0.357	0.342
Within R^2	0.000	0.000	0.000	0.000
Observations	27,824,414	24,336,456	22,370,058	19,736,823

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: OECD countries, more balanced panel where observations appear at least 80% of the periods after they first enter the sample. CIT refers to the central CIT.

TABLE 1.A.8: CORPORATE INCOME TAX RATES AND REPORTING GAPS: EACH SIDE SEPARATELY, ALTERNATIVE GAP MEASURE

	(1) Value gap, log	(2) Quantity gap, log	(3) Value gap, log	(4) Quantity gap, log
CIT Exporter	-0.00105***	-0.00168***		
	(-4.01)	(-5.23)		
CIT Importer			0.000581	0.000611
			(1.38)	(1.44)
Tariff	0.000680***	-0.000538**	-0.000269	-0.000542
	(4.74)	(-2.48)	(-1.02)	(-1.49)
Importer-export-product FE	✓	✓	✓	✓
Importer-year FE	✓	✓		
Exporter-year FE			✓	✓
Exporter-specific time trend	✓	✓		
Importer-specific time trend			✓	✓
Controls	✓	✓	✓	✓
Adjusted R^2	0.327	0.325	0.339	0.335
Within R^2	0.000	0.000	0.000	0.000
Observations	44,623,148	39,218,517	33,866,670	30,118,072

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: OECD countries. CIT refers to the central CIT. The reporting gaps are measured using expressions (1.12).

Table 1.A.9: corporate income tax rates and reporting gaps: Each side separately, net-of-tax rate

	(1) Value gap, log	(2) Quantity gap, log	(3) Value gap, log	(4) Quantity gap, log
CIT Exporter, log(net-of-tax rate)	0.103*** (3.79)	0.166*** (4.52)		
CIT Importer, log(net-of-tax rate)			-0.0582 (-1.40)	-0.00623 (-0.14)
Tariff	0.000964*** (4.10)	-0.00135*** (-3.19)	-0.000651 (-1.59)	-0.00108 (-1.62)
Importer-export-product FE	✓	✓	✓	✓
Importer-year FE	✓	✓		
Exporter-year FE			✓	✓
Exporter-specific time trend	✓	✓		
Importer-specific time trend			✓	✓
Controls	✓	✓	✓	✓
Adjusted R^2	0.326	0.322	0.335	0.329
Within R^2	0.000	0.000	0.000	0.000
Observations	44,623,145	38,666,718	33,866,668	29,732,923

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: OECD countries. CIT refers to the central CIT. CIT rates are specified as $\log(1 - \text{CIT})$ in the regressions, i.e. as log of the net-of-tax rate.

Table 1.A.10: corporate income tax rates and reporting gaps: World sample

	(1) Value gap, log	(2) Quantity gap, log	(3) Value gap, log	(4) Quantity gap, log
CIT Exporter	-0.00139*** (-3.33)	-0.00270*** (-4.98)		
CIT Importer			0.000427 (0.75)	-0.000363 (-0.55)
Tariff	0.000972*** (4.37)	-0.000941*** (-2.61)	0.000012 (0.04)	-0.00051 (-1.28)
Importer-export-product FE	✓	✓	✓	✓
Importer-year FE	✓	✓		
Exporter-year FE			✓	✓
Exporter-specific time trend	✓	✓		
Importer-specific time trend			✓	✓
Controls	✓	✓	✓	✓
Adjusted R^2	0.345	0.339	0.353	0.346
Within R^2	0.000	0.000	0.000	0.000
Observations	55,100,143	47,906,134	49,167,239	42,841,531

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: world sample. CIT refers to the combined CIT.

TABLE 1.A.11: CORPORATE INCOME TAX RATES AND REPORTING GAPS: WORLD SAMPLE WITHOUT TIME TRENDS

	(1) Value gap, log	(2) Quantity gap, log	(3) Value gap, log	(4) Quantity gap, log
CIT Exporter	-0.00185***	0.000392		
	(-3.70)	(0.56)		
CIT Importer			0.000963*	0.000444
			(1.70)	(0.59)
Tariff	0.000996***	-0.000672*	0.000614*	-0.000962**
	(4.29)	(-1.83)	(1.76)	(-1.97)
Importer-export-product FE	✓	✓	✓	✓
Importer-year FE	✓	✓		
Exporter-year FE			✓	✓
Exporter-specific time trend				
Importer-specific time trend				
Controls	✓	✓	✓	✓
Adjusted R^2	0.344	0.338	0.352	0.345
Within R^2	0.000	0.000	0.000	0.000
Observations	55,100,143	47,906,134	49,167,239	42,841,531

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: world sample. CIT refers to the combined CIT.

TABLE 1.A.12: CORPORATE INCOME TAX RATES AND REPORTING GAPS: WITHIN-EU TRADE WITHOUT TIME TRENDS

	(1) Value gap, log	(2) Quantity gap, log	(3) Value gap, log	(4) Quantity gap, log
CIT Exporter	-0.00221***	-0.00125*		
	(-4.00)	(-1.91)		
Log(Intrastat threshold), Exporter	-0.0146	-0.0207		
	(-1.41)	(-1.60)		
Standard VAT rate, Exporter	0.00965***	0.00854***		
	(3.53)	(2.73)		
CIT Importer			0.00282***	0.00239***
			(4.14)	(3.10)
Log(Intrastat threshold), Importer			0.0388***	0.0443***
			(4.15)	(4.37)
Standard VAT rate, Importer			-0.00878***	-0.00962***
			(-3.37)	(-2.94)
Importer-export-product FE	✓	✓	✓	✓
Importer-year FE	✓	✓		
Exporter-year FE			✓	✓
Exporter-specific time trend				
Importer-specific time trend				
Controls	✓	✓	✓	✓
Adjusted R^2	0.316	0.309	0.317	0.308
Within R^2	0.000	0.000	0.000	0.000
Observations	16,899,957	15,167,030	15,891,282	14,271,876

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: OECD countries. CIT refers to the central CIT.

TABLE 1.A.13: CORPORATE INCOME TAX RATES AND REPORTING GAPS: WITHIN-EU TRADE VERSUS NOT

	(1) Value gap, log	(2) Quantity gap, log	(3) Value gap, log	(4) Quantity gap, log
Within-EU=1	-0.0765*	-0.0674	-0.0737	-0.0362
	(-1.68)	(-1.25)	(-1.42)	(-0.58)
CIT Importer	-0.00193**	-0.00278***	-0.00218**	-0.00160
	(-2.33)	(-2.75)	(-2.28)	(-1.28)
Within-EU=1 × CIT Importer	0.00447***	0.00425***	0.00488***	0.00357**
	(4.43)	(3.41)	(4.23)	(2.29)
Importer-exporter-product FE	✓	✓	✓	✓
Exporter-year FE	✓	✓	✓	✓
Importer-specific time trend	✓	✓		
Controls	✓	✓	✓	✓
Adjusted R^2	0.334	0.328	0.335	0.329
Within R^2	0.000	0.000	0.000	0.000
Observations	33,866,668	29,732,923	33,866,668	29,732,923

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: OECD countries. CIT refers to the central CIT.

TABLE 1.A.14: DIMENSIONS OF CHANGE OF CORPORATE TAX BASE

	Dimension of base change
1	The research and development tax credit (R&D credit)
2	Credits for foreign taxes paid
3	The tax treatment of foreign companies, such as credits or other incentives to attract foreign investment
4	Policies that target evasion or avoidance by companies
5	Investment credits or other tax incentives to promote investment
6	Accelerated depreciation or other depreciation allowances
7	Other tax rates that may affect the corporate tax base, e.g., net worth tax on corporations or extraordinary profits tax
8	Loss carry-forward rules
9	Loss carry-back rules
10	Thin capitalization rules
11	Controlled foreign company (CFC) legislation
12	Other changes to the corporate tax base

Note: Dimensions along which the corporate income tax base can change. Source: Kawano and Slemrod (2016, p.4).

TABLE 1.A.15: DESCRIPTIVE STATISTICS: CHANGES IN CORPORATE INCOME TAX BASE

	N	Mean	Std	Min	25%	50%	75%	Max
Δ evasion crackdown	750	.14	.35	0	0	0	0	1
Cum. Δ evasion crackdown	750	1.9	2.2	0	0	1	3	12
Δ tax base	750	.097	.91	-4	0	0	0	4
Cum. Δ tax base	750	1.3	4.1	-12	-1	1	3	15

Note: descriptive statistics of changes and cumulative changes in corporate income tax base. The first two lines refer to changes in policies fighting evasion, whereas the two last lines aggregate changes along all dimensions listed in table 1.A.14. Source: Kawano and Slemrod (2016).

TABLE 1.A.16: STATUTORY CIT RATE VERSUS EFFECTIVE AVERAGE TAX RATE — WITHIN EU TRADE

	Dep. variable: Value gap, in log			
	(1)	(2)	(3)	(4)
EATR Exporter	-0.00874***			
	(-7.71)			
CIT Exporter		-0.00400***		
		(-5.14)		
Log(Intrastat threshold), exporter	-0.0284**	-0.0230*		
	(-2.50)	(-1.95)		
Standard VAT rate, exporter	0.00865***	0.00961***		
	(3.23)	(3.54)		
EATR Importer			0.00766***	
			(6.12)	
CIT Importer				0.00385***
				(4.79)
Log(Intrastat threshold), importer			0.0368***	0.0465***
			(3.54)	(4.74)
Standard VAT rate, importer			-0.00910***	-0.00887***
			(-3.51)	(-3.49)
Importer-export-product FE	✓	✓	✓	✓
Importer-year FE	✓	✓		
Exporter-year FE			✓	✓
Controls	✓	✓	✓	✓
Adjusted R^2	0.330	0.329	0.329	0.331
Within R^2	0.000	0.000	0.000	0.000
Observations	15,919,782	15,675,922	15,626,162	14,667,247

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: EU countries, 1998-2017. CIT refers to the central CIT. EATR is the effective average tax rate at the corporation level as defined and computed by Spengel et al. (2019).

TABLE 1.A.17: CIT AND REPORTING GAPS: TRADE IN SERVICES (EACH SIDE SEPARATELY)

	Dep. variable: Value gap					
	OECD sample					
	(1) All flows	(2) All flows	(3) 1-digit	(4) 1-digit	(5) 2-digit	(6) 2-digit
CIT Exporter	-0.0107***		-0.00544		-0.0153***	
	(-2.97)		(-1.40)		(-3.33)	
CIT Importer		0.00380		0.00353		0.00312
		(1.22)		(0.98)		(0.73)
Adjusted R^2	0.613	0.608	0.576	0.555	0.631	0.628
Within R^2	0.003	0.001	0.004	0.001	0.004	0.003
Observations	117,982	118,706	37,321	37,176	36,343	36,636
	within-EU sample					
CIT Exporter	-0.0168***		-0.0135***		-0.0181***	
	(-4.09)		(-2.98)		(-3.69)	
CIT Importer		0.00675*		0.00728*		0.00579
		(1.95)		(1.73)		(1.31)
Adjusted R^2	0.625	0.627	0.590	0.589	0.646	0.646
Within R^2	0.002	0.001	0.003	0.002	0.003	0.002
Observations	106,612	106,557	30,713	30,668	34,378	34,264
Year FE	✓	✓	✓	✓	✓	✓
Importer-exporter-product FE	✓	✓	✓	✓	✓	✓
Importer-year FE	✓		✓		✓	
Exporter-year FE		✓		✓		✓
Controls	✓	✓	✓	✓	✓	✓

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: OECD and EU countries. CIT refers to the central CIT. The columns refer to different levels of aggregation: (1)-(2) all flows; (3)-(4) 1-digit flows (12 categories); (5)-(6) 2-digit flows (32 categories).

TABLE 1.A.18: TARIFFS AND REPORTING GAPS

	(1) Value gap, log	(2) Quantity gap, log
Tariff	0.00164***	-0.000223
	(8.33)	(-0.69)
Importer-exporter-product FE	✓	✓
Importer-exporter-year FE	✓	✓
Adjusted R^2	0.346	0.340
Within R^2	0.000	0.000
Observations	60,495,458	52,614,533

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: world sample.

TABLE 1.A.19: TARIFFS AND DIFFERENTIATED PRODUCTS

	(1) Value gap, log	(2) Quantity gap, log	(3) Value gap, log	(4) Quantity gap, log
Tariff	0.00167***	0.00137***	0.00179***	-0.0000874
	(5.10)	(4.98)	(8.18)	(-0.20)
Differentiated × Tariff	0.00265***	-0.0000799	-0.000234	-0.000217
	(7.70)	(-0.19)	(-0.82)	(-0.46)
Product FE	✓	✓		
Importer-exporter-product FE			✓	✓
Importer-exporter-year FE	✓	✓	✓	✓
Adjusted R^2	0.057	0.056	0.346	0.340
Within R^2	0.000	0.000	0.000	0.000
Observations	62,978,078	54,965,946	60,495,458	52,614,533

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: world sample. Differentiated is a dummy variable that takes value 1 if the product is differentiated, according to the liberal definition of Rauch (1999). This dummy is absorbed by the product fixed effect when non-interacted with tariffs.

TABLE 1.A.20: TARIFF EVASION AND CORRUPTION LEVELS

	(1) Value gap, log	(2) Quantity gap, log	(3) Value gap, log	(4) Quantity gap, log
Tariff	0.00130***	-0.00136**	0.00140***	-0.000586
	(5.03)	(-2.55)	(6.37)	(-1.51)
High corruption, importer × Tariff	0.000496**	0.00182***		
	(2.09)	(4.01)		
High corruption, both × Tariff			0.000800**	0.000882**
			(2.34)	(2.00)
Importer-exporter-product FE	✓	✓	✓	✓
Importer-exporter-year FE	✓	✓	✓	✓
Adjusted R^2	0.344	0.340	0.344	0.339
Within R^2	0.000	0.000	0.000	0.000
Observations	58,165,200	50,635,680	57,278,841	49,863,946

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: world sample. High corruption is a dummy that takes value 1 if a country has a corruption index below 3, using the index from the ICRG.

TABLE 1.A.21: TARIFF EVASION THROUGH PRODUCT MISCLASSIFICATION

	(1) Value gap, log	(2) Value gap, log	(3) Quantity gap, log	(4) Quantity gap, log
Tariff	0.00297***	0.00475***	0.00104***	0.00306***
	(9.69)	(10.98)	(2.94)	(5.45)
Tariff on similar products, HS4	-0.00209***		-0.00281***	
	(-7.69)		(-7.81)	
Tariff on similar products, HS5		-0.00445***		-0.00572***
		(-12.11)		(-9.95)
Importer-exporter-product FE	✓	✓	✓	✓
Importer-exporter-year FE	✓	✓	✓	✓
Adjusted R^2	0.357	0.376	0.351	0.366
Within R^2	0.000	0.000	0.000	0.000
Observations	49,340,679	20,297,522	42,613,817	17,926,066

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: world sample.

TABLE 1.A.22: EFFECTS OF TARIFFS ON COMPONENTS OF THE REPORTING GAPS

	(1) Imports value^i, log	(2) Exports value^e, log	(3) Imports quantity^i, log	(4) Exports quantity^e, log
Tariff	-0.00355***	-0.00170***	-0.00321***	-0.00375***
	(-10.31)	(-5.53)	(-7.24)	(-9.06)
Importer-exporter-product FE	✓	✓	✓	✓
Importer-exporter-year FE	✓	✓	✓	✓
Adjusted R^2	0.755	0.749	0.785	0.788
Within R^2	0.000	0.000	0.000	0.000
Observations	61,805,551	61,805,551	56,610,573	56,772,659

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: world sample. The dependent variables are the log of, in order of the columns: reported imports and exports, reported imported and exported quantities. The superscript indicates the reporting country.

TABLE 1.A.23: NON-LINEAR RELATIONSHIP BETWEEN TARIFFS AND REPORTING GAPS

	(1) Value gap, log	(2) Quantity gap, log
Tariff	0.00193***	-0.000312
	(8.59)	(-0.82)
Tariff squared	-0.000000911***	0.000000262
	(-5.41)	(1.27)
Importer-exporter-product FE	✓	✓
Importer-exporter-year FE	✓	✓
Adjusted R^2	0.346	0.340
Within R^2	0.000	0.000
Observations	60,495,458	52,614,533

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: world sample.

TABLE 1.A.24: COMPARISON TO OTHER STUDIES

Paper	Setup	Estimated α
Fisman and Wei (2004, table 5)	China-HK, 1998, cross-section	3%
Mishra et al. (2008, table 2)	India and partners, 1990s, panel	0.1%
Javorcik and Narciso (2008, table 3)	Germany and Eastern European Countries, 1992-2003, panel	0.9% - 4.5%
Stoyanov (2012, table 3)	Canada-US, 1989, cross-section	0.85% - 4.6%
Javorcik and Narciso (2017, table 14)	15 countries accessing the WTO in 1996-2008, panel	0.73% - 1.31%

TABLE 1.A.25: TARIFF EVASION AND WTO MEMBERSHIP

	(1) Value gap, log	(2) Quantity gap, log	(3) Value gap, log	(4) Quantity gap, log
Tariff	0.00190***	-0.00258*	0.00184**	-0.00458**
	(3.28)	(-1.88)	(2.50)	(-2.10)
WTO Importer=1 × Tariff	-0.000353	0.00237*	-0.000408	0.00360*
	(-0.61)	(1.74)	(-0.55)	(1.65)
Importer-exporter-product FE	✓	✓	✓	✓
Importer-exporter-year FE	✓	✓	✓	✓
Adjusted R^2	0.346	0.341	0.352	0.351
Within R^2	0.000	0.000	0.000	0.000
Observations	58,574,541	50,929,118	43,183,847	36,239,930

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: world sample, restricted to observations where the exporting country is in the WTO. The WTO membership dummy is not included alone in the regression, as it would be subsumed by the importer-exporter-year fixed effect. Columns (1) and (2) are based on the whole sample, whereas (3) and (4) are estimated solely on differentiated products.

TABLE 1.A.26: TARIFF AND REPORTING GAPS WITH DIFFERENT FIXED EFFECTS

	Dependent variable: Value gap, log				
	(1)	(2)	(3)	(4)	(5)
Tariff	0.00341***	0.00221***	0.00251***	0.000661*	0.000395*
	(8.04)	(7.50)	(7.35)	(1.81)	(1.72)
Constant	0.0152**				
	(2.12)				
Adjusted R^2	0.000	0.046	0.051	0.337	0.338
Within R^2		0.000	0.000	0.000	0.000
Observations	62,996,855	62,995,490	62,995,440	60,517,075	60,517,075

	Dependent variable: Quantity gap, log				
	(1)	(2)	(3)	(4)	(5)
Tariff	0.000389	0.000951***	0.00118***	-0.000579	-0.000616*
	(1.05)	(4.49)	(5.81)	(-1.14)	(-1.82)
Constant	0.0132*				
	(1.67)				
Adjusted R^2	0.000	0.041	0.051	0.328	0.330
Within R^2		0.000	0.000	0.000	0.000
Observations	54,984,914	54,983,559	54,983,504	52,636,106	52,636,106
Year FE		✓	✓	✓	✓
Importer-exporter FE		✓	✓		
Product FE		✓	✓		
Importer-year FE			✓		
Exporter-year FE			✓	✓	✓
Importer-exporter-product FE				✓	✓
Importer-specific time trend					✓

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: world sample.

TABLE 1.A.27: TARIFF EVASION OVER TIME

	(1) Value gap, log	(2) Quantity gap, log
1988-1993 × Tariff	-0.000341	-0.00683***
	(-0.83)	(-6.42)
1994-1998 × Tariff	0.00158***	-0.00246***
	(4.79)	(-3.61)
1999-2003 × Tariff	0.00201***	0.00143***
	(7.95)	(4.29)
2004-2008 × Tariff	0.00206***	0.000980***
	(8.83)	(3.64)
2009-2013 × Tariff	0.00183***	0.00122***
	(7.80)	(4.33)
2014-2017 × Tariff	0.00194***	0.00104***
	(6.66)	(3.17)
Importer-exporter-product FE	✓	✓
Importer-exporter-year FE	✓	✓
Adjusted R^2	0.346	0.340
Within R^2	0.000	0.000
Observations	60,495,458	52,614,533

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: world sample. Tariff is interacted with time period dummies for periods $1988-1993, \ldots, 2014-2017$.

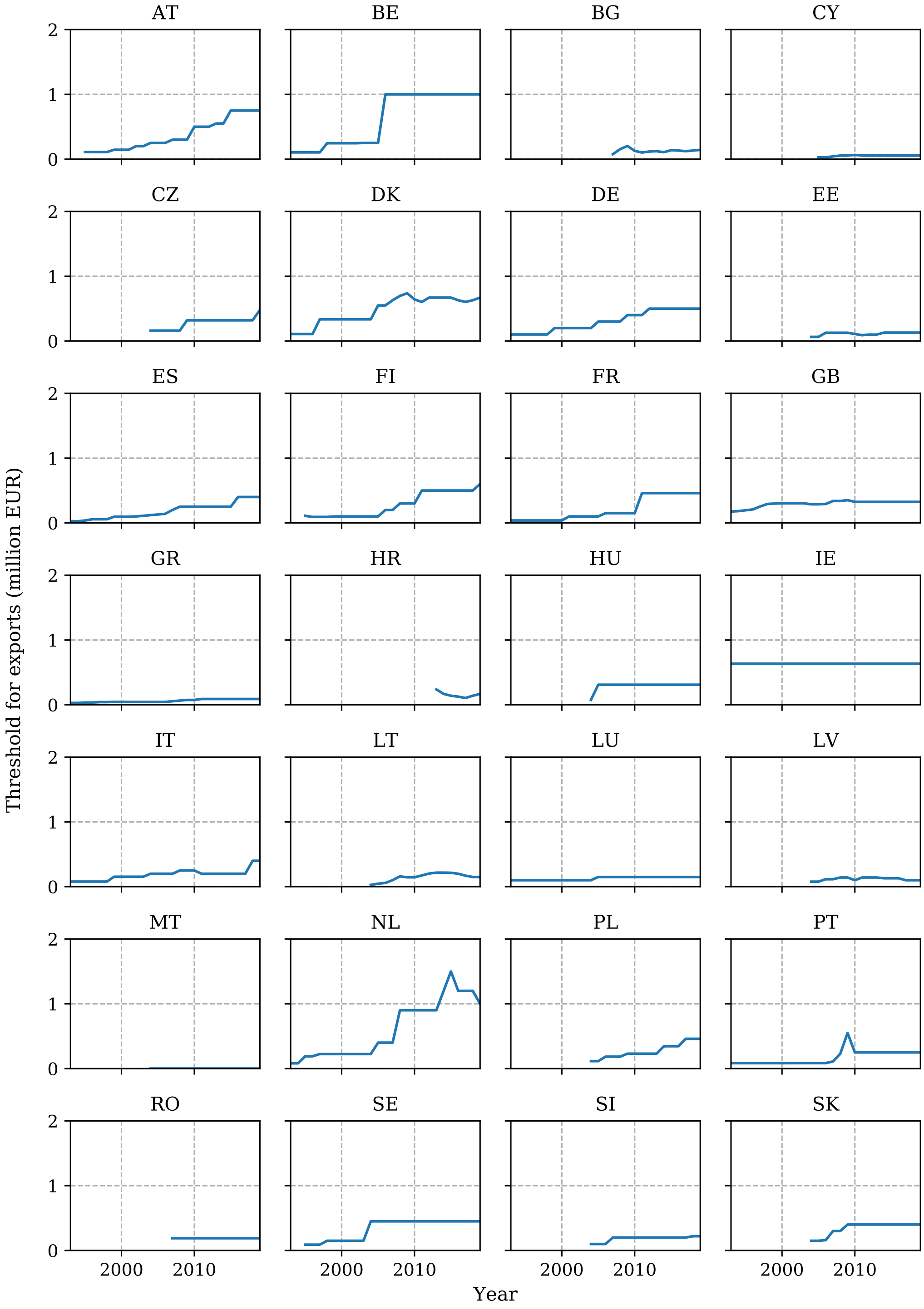
TABLE 1.A.28: TARIFF EVASION ACROSS INCOME LEVELS

	(1) Value gap, log	(2) Quantity gap, log
Low × Tariff	0.00291***	0.00207***
	(7.00)	(3.84)
Lower-middle × Tariff	0.00244***	0.000330
	(6.33)	(0.49)
Upper-middle × Tariff	0.00144***	-0.000118
	(5.35)	(-0.38)
High × Tariff	-0.00199***	-0.00488***
	(-3.75)	(-4.66)
Importer-exporter-product FE	✓	✓
Importer-exporter-year FE	✓	✓
Adjusted R^2	0.346	0.340
Within R^2	0.000	0.000
Observations	60,489,384	52,609,632

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Data: world sample. Income group are as defined by the World Bank. To avoid countries switching income group over the sample periods, groups are assigned as of 2000.

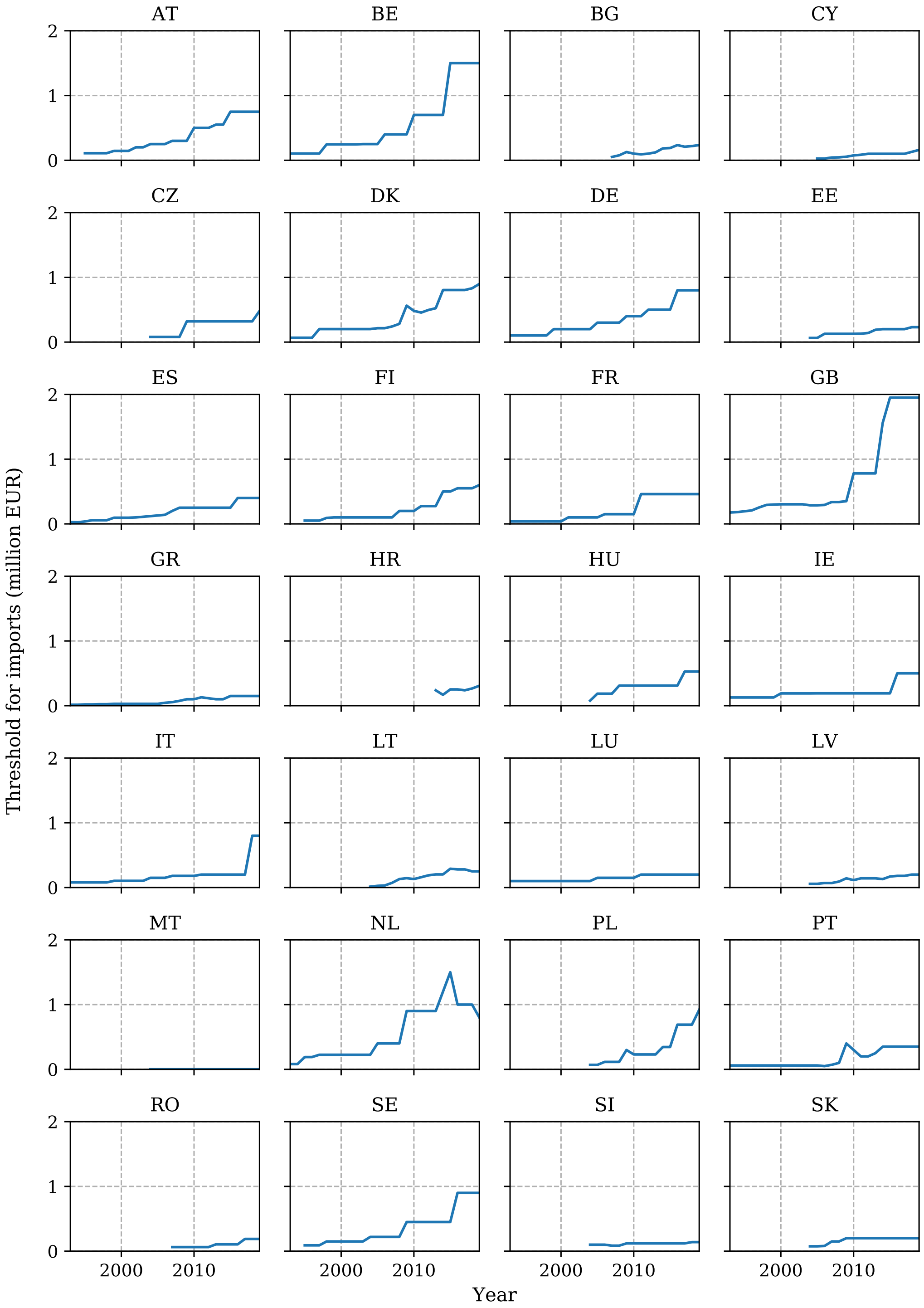
Appendix 1.B Additional figures

FIGURE 1.B.1: INTRASTAT EXEMPTION THRESHOLD FOR EXPORTS, BY EU COUNTRY



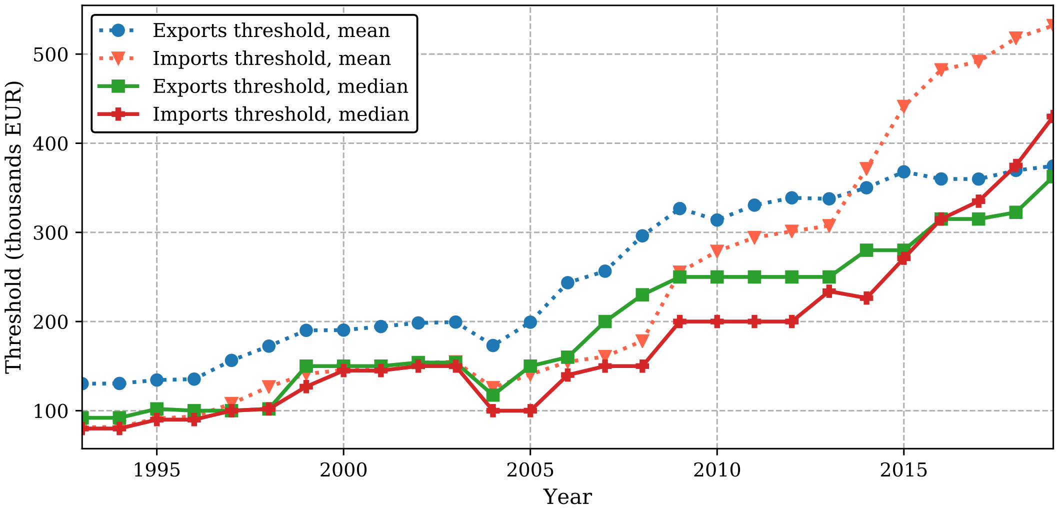
Note: Evolution of threshold for exports for each EU member over time. Source: own collection.

FIGURE 1.B.2: INTRASTAT EXEMPTION THRESHOLD FOR IMPORTS, BY EU COUNTRY



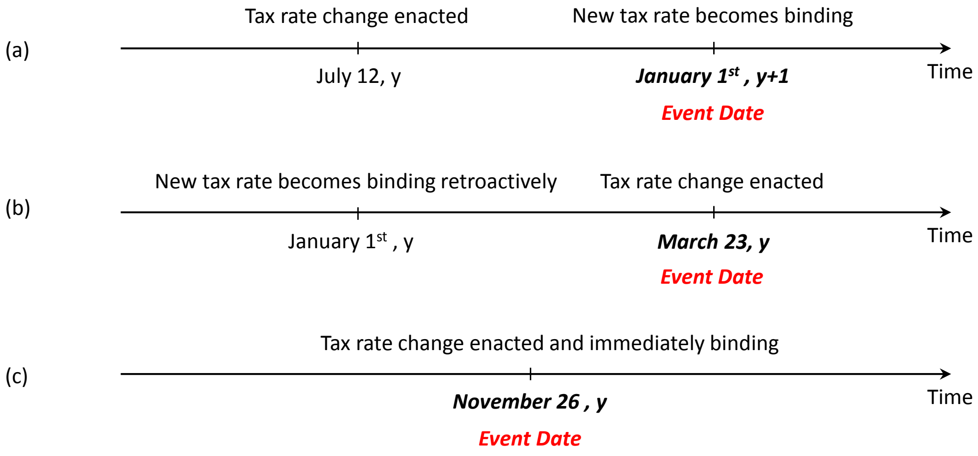
Note: Evolution of threshold for imports for each EU member over time. Source: own collection.

Figure 1.B.3: Mean and median exemption thresholds for imports and exports over time



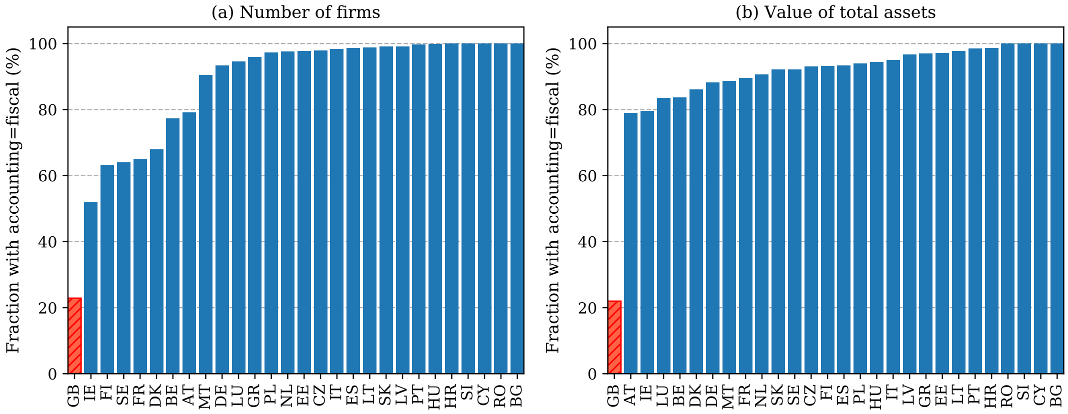
Note: Average and median threshold for imports and exports across EU countries over time. Source: own collection.

Figure 1.B.4: Definition of event date for corporate tax changes



Note: Definition of the event date for corporate income tax rate changes. The event date is defined as the latest of the date at which the new tax rate became binding and the date at which firms became aware of it.

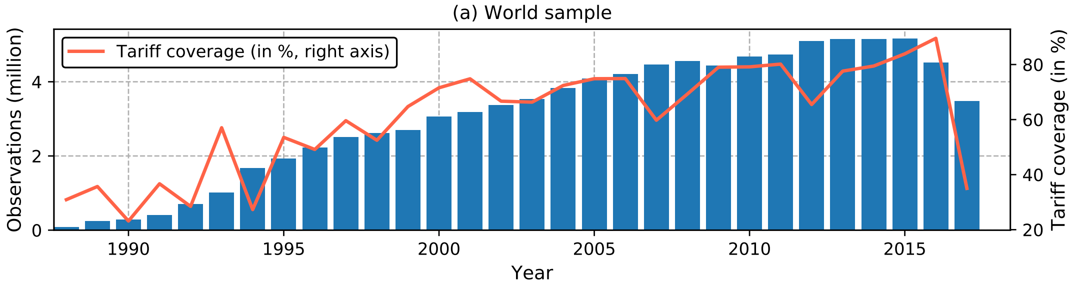
FIGURE 1.B.5: ACCOUNTING PERIOD AND FISCAL YEAR IN THE EU

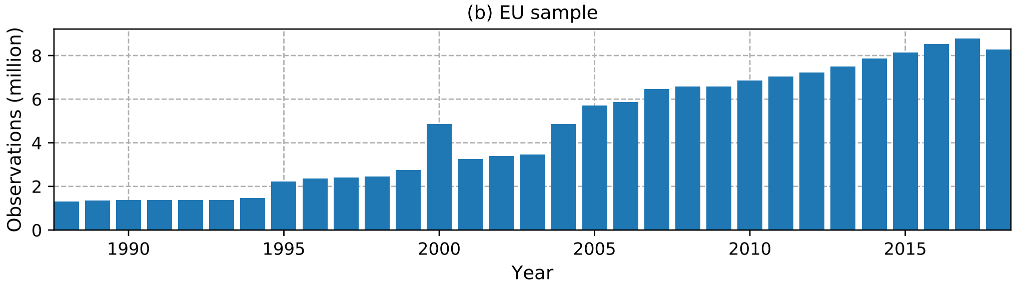


(a) Number of firms (b) Value of total assets

Note: The sample is Amadeus over the years 2015-2017. This figure plots (a) the average fraction (across years) of firms whose accounting period coincides with the fiscal year in their country of incorporation; (b) the same as in (a), where firms are weighted by the value of their assets. The UK's — in hatched orange — fiscal year does not coincide with the calendar year, but starts on April 1st. The average number of firms by year for each country can be found in table 1.A.3.

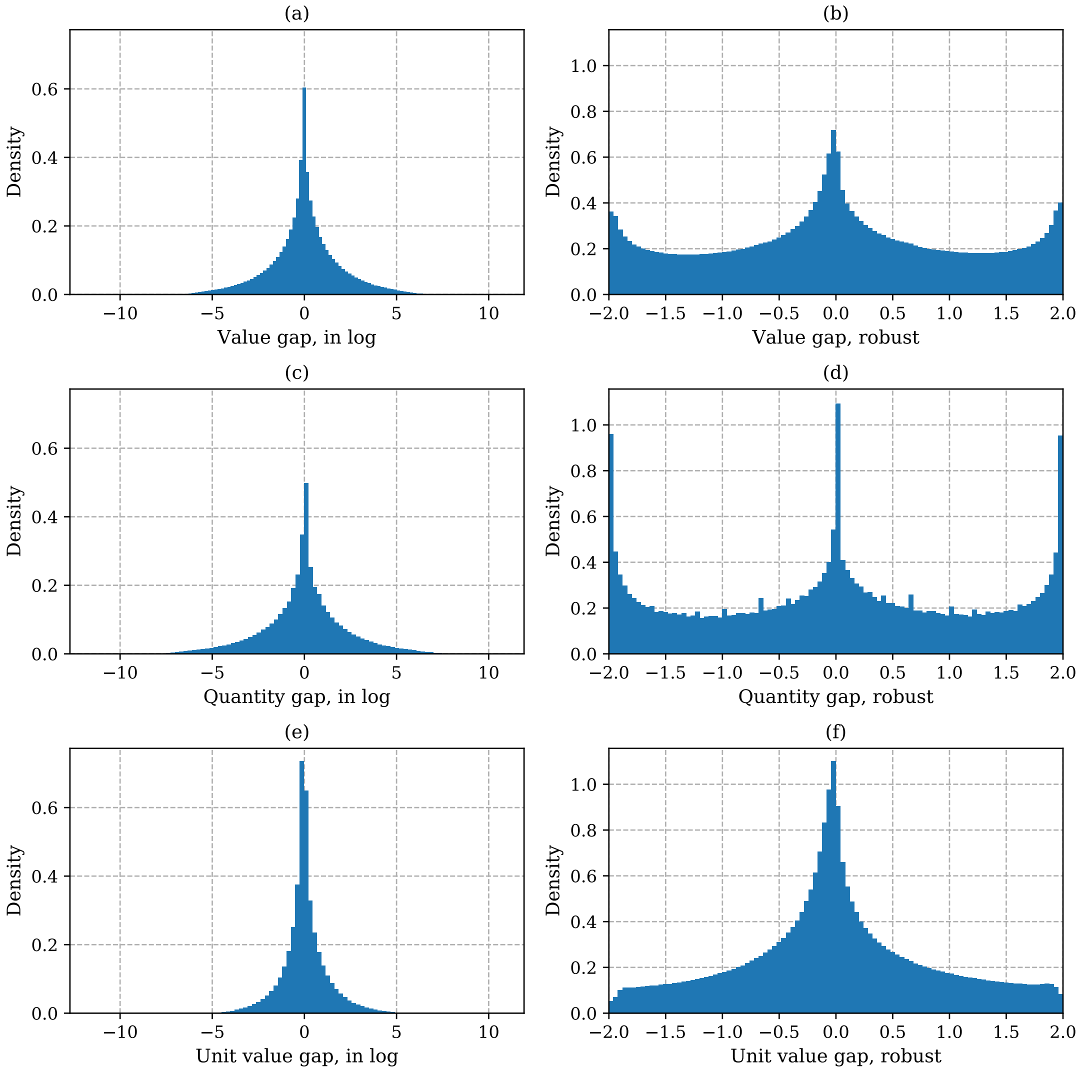
FIGURE 1.B.6: NUMBER OF OBSERVATION PER YEAR


(a) World sample


(b) EU sample

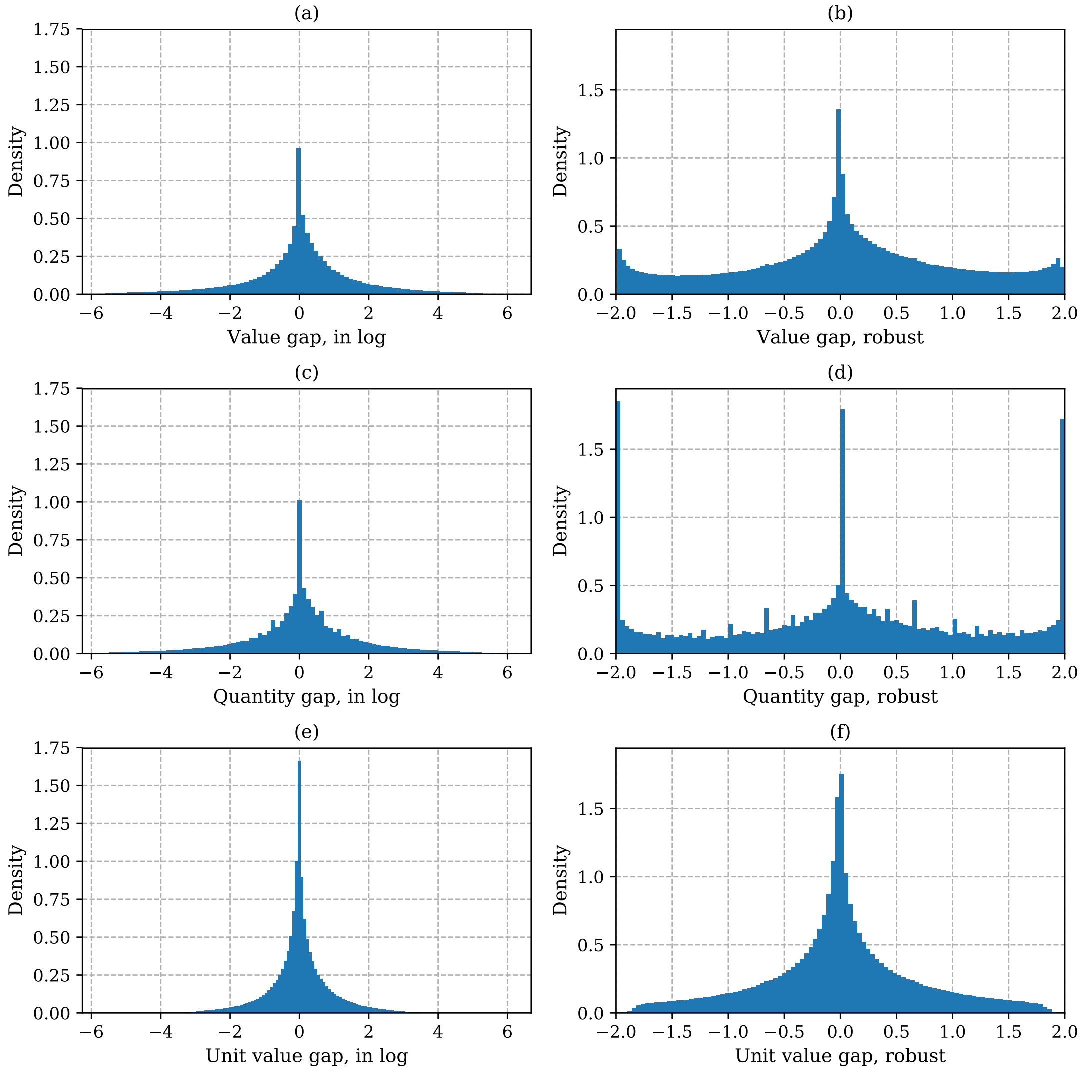
Note: This figure plots the number of observations in the sample for which the reporting gap in terms of value is available. Panel (a) plots this data for the world sample, and also displays the share of observations for which tariff information is available. Panel (b) plots this data for the within-EU sample, which is larger as observations are available at a monthly frequency. The data is at the HS 6-digit level in both figures.

FIGURE 1.B.7: HISTOGRAMS OF REPORTING GAPS IN THE WORLD SAMPLE



Note: histograms for each measure of the reporting gaps, in the world sample. The first column plots the gaps in logs, and the second column using the robustness measure as defined in equation (1.12). The top and bottom percentiles in terms of the gaps by reporting country are dropped from the sample.

Figure 1.B.8: Histograms of reporting gaps in the within-EU sample

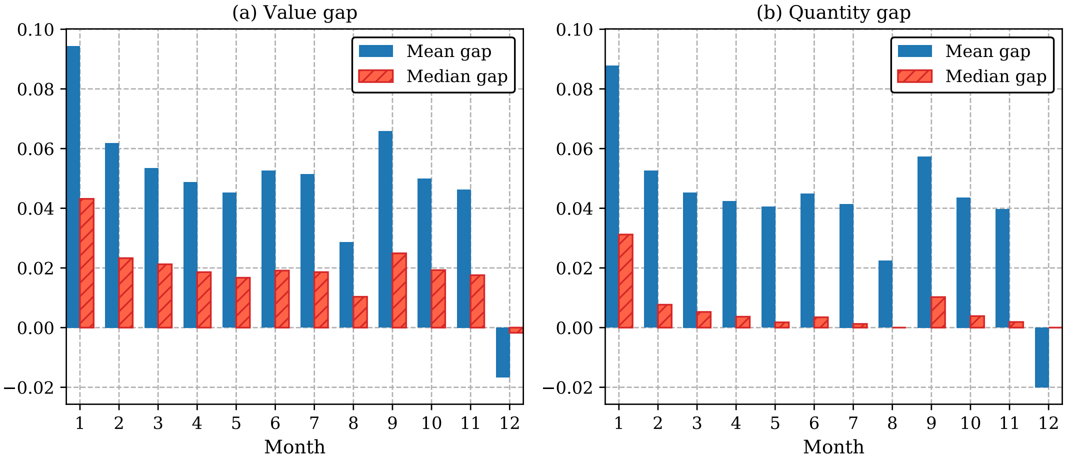


Note: histograms for each measure of the reporting gaps, in the within-EU sample. The first column plots the gaps in logs, and the second column using the robustness measure as defined in equation (1.12). The top and bottom percentiles in terms of the gaps by reporting country are dropped from the sample.

FIGURE 1.B.9: AGGREGATE TRADE FLOWS AND REPORTING GAPS IN THE WORLD SAMPLE

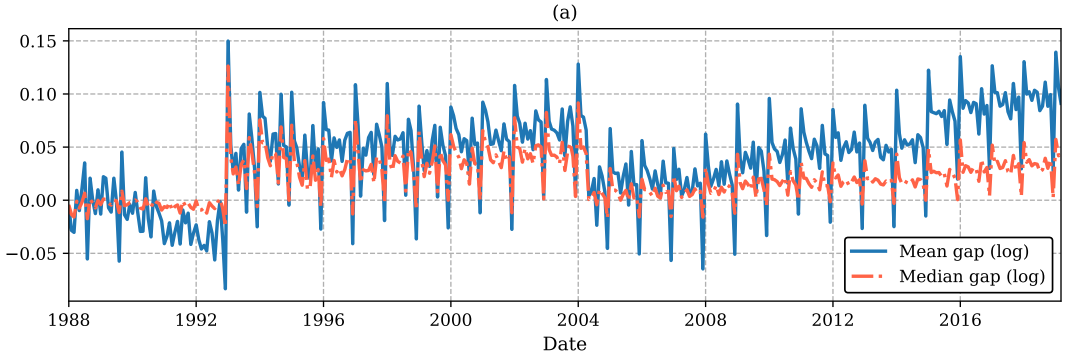
Note: aggregated trade flows (from HS 6-digit level products) for which both the importer and exporter reports are available, and the resulting aggregate reporting gap. In case products come from several HS versions for a given reporter-year tuple, only the most recent version is used for the aggregation. The trade gap is measured as in equation (1.12), so a negative number means that imports are higher than exports.

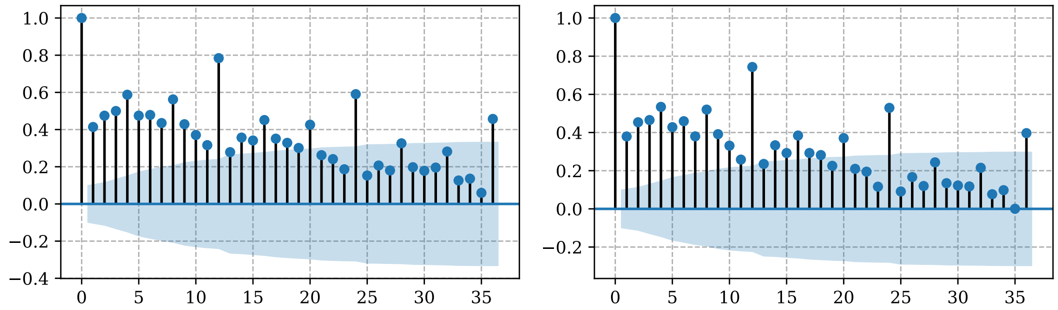
FIGURE 1.B.10: MEAN AND MEDIAN REPORTING GAPS OVER MONTHS



Note: Mean and median gaps in log over months in the within-EU sample, across all products, country pairs and years.

FIGURE 1.B.11: TIME SERIES PROPERTIES OF REPORTING GAPS



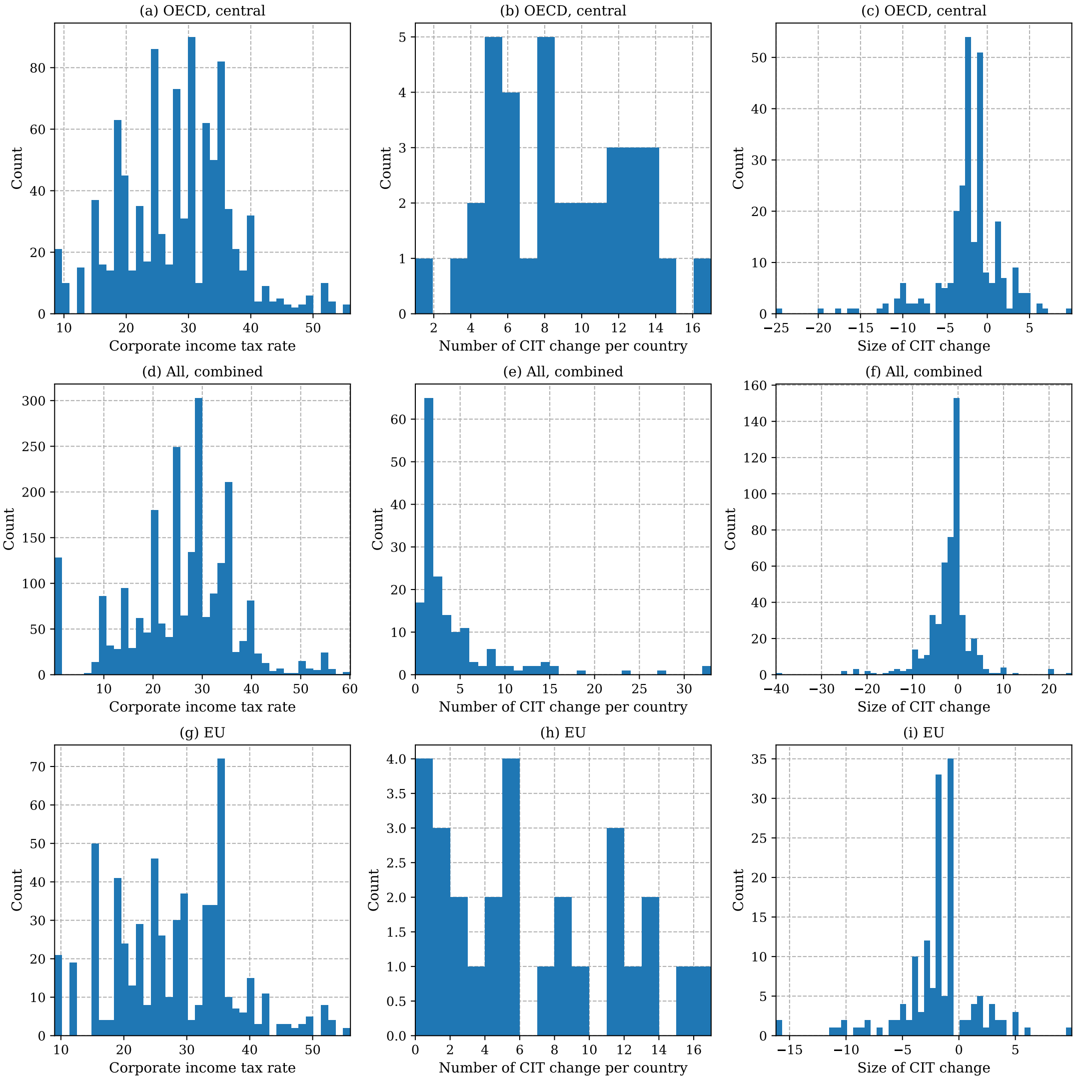


(b) Autocorrelation of mean value gap

(c) Autocorrelation of median value gap

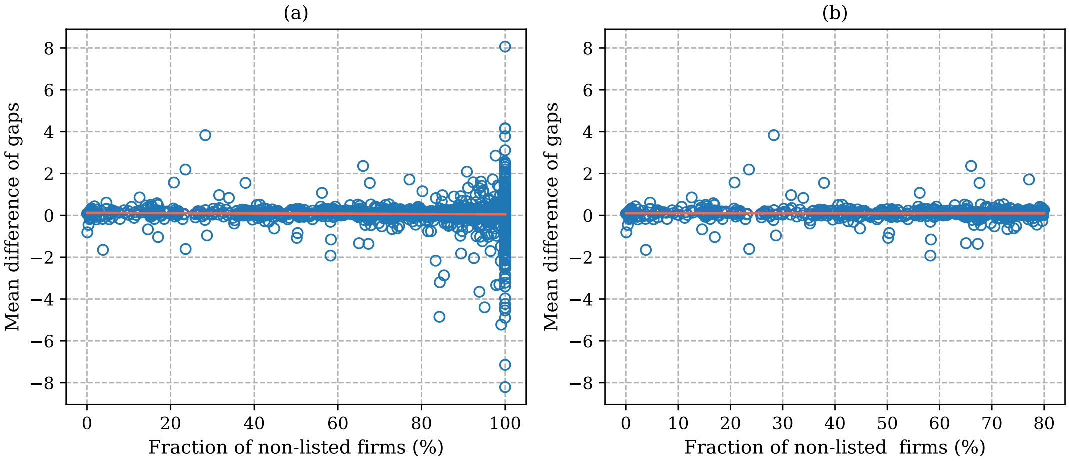
Note: Panel (a) displays the mean and median value gap over the whole sample period. Panel (b) and (c) show the estimated auto-correlation functions for the series in panel (a), at lags of up to 36 months. The shaded areas indicate the 95% confidence interval bands around zero. In panel (a), the first structural break coincides with the creation of the European single market in January 1993. The second break coincides with the extension of the EU from 10 to 25 members in 2004.

FIGURE 1.B.12: HISTOGRAMS OF CORPORATE INCOME TAX RATES



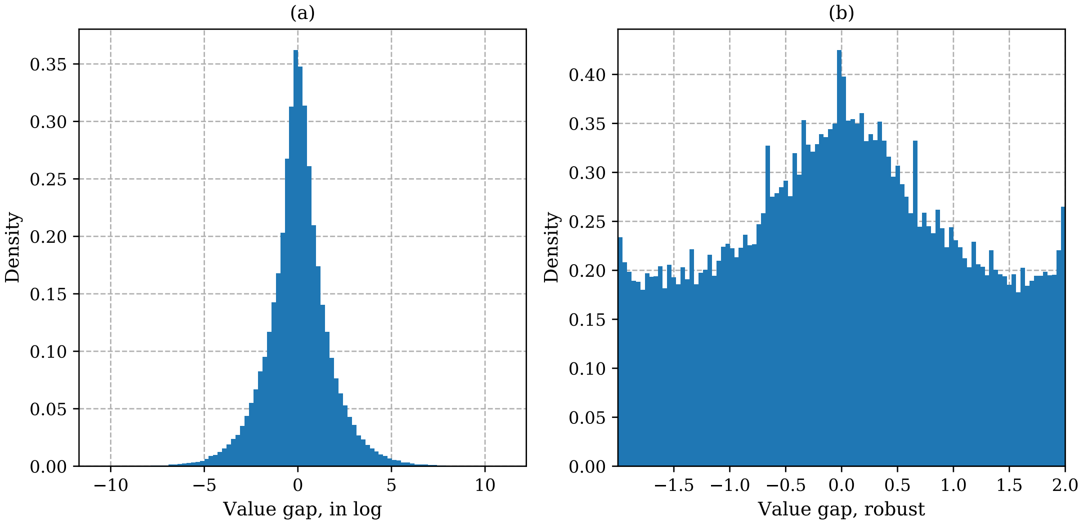
Note: Panels (a)-(c): all graphs based on central CIT rate in OECD countries. Panels (d)-(f): all graphs based on combined CIT for all countries covered by the OECD and KPMG data. Panels (g)-(i): all graphs based on the within-EU sample. The first column contains histograms of the CIT values; the second column contains histograms of the number of tax change within countries; the third column contains histograms of the size of the tax changes.

FIGURE 1.B.13: DECEMBER-JANUARY GAP DIFFERENCE: LISTED VERSUS NON-LISTED FIRMS



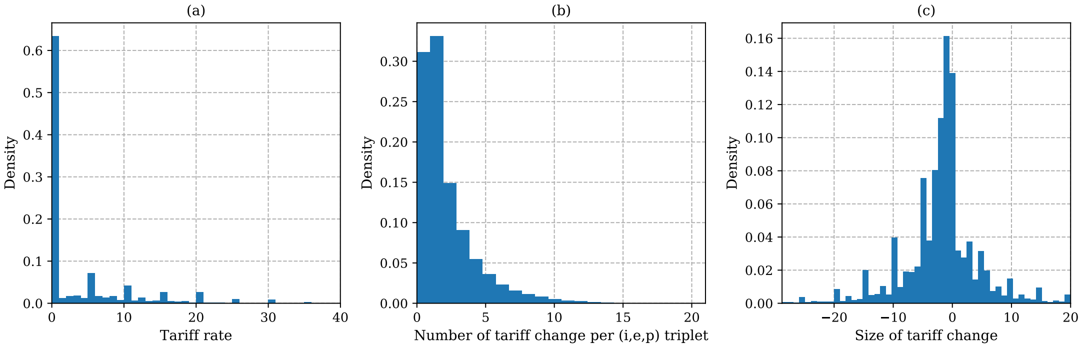
Note: Scatter plots of mean difference in the reporting gaps between December and January against the fraction of firms that are not listed (in a given year, weighted by their total assets). Each dot represents a country-sector-year triple, where sectors are defined according to the NACE revision 2 classification at the 2-digit level. Gaps and fractions of firms are matched on the exporting country. Panel (a) does not impose any restrictions on the data, whereas panel (b) only shows data where the fraction of non-listed firms is lower than 80%. The data on listing status is from BVD Amadeus in years 2015, 2016 and 2017.

FIGURE 1.B.14: HISTOGRAMS OF THE REPORTING GAPS: TRADE IN SERVICES



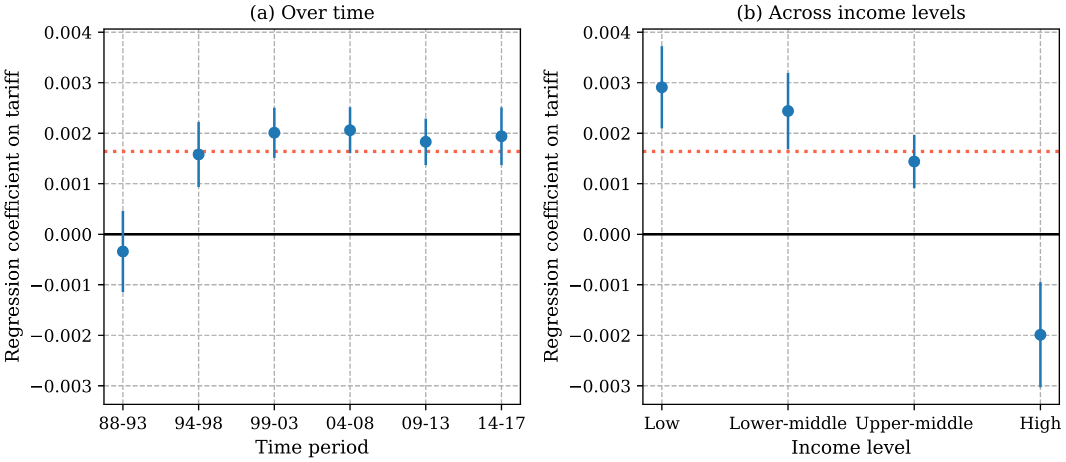
Note: histograms of the reporting gaps in the sample of trade in services. Panel (a) and (b) show the histograms of the gaps measured according to expressions (1.11) and (1.12), respectively.

FIGURE 1.B.15: HISTOGRAMS OF TARIFFS



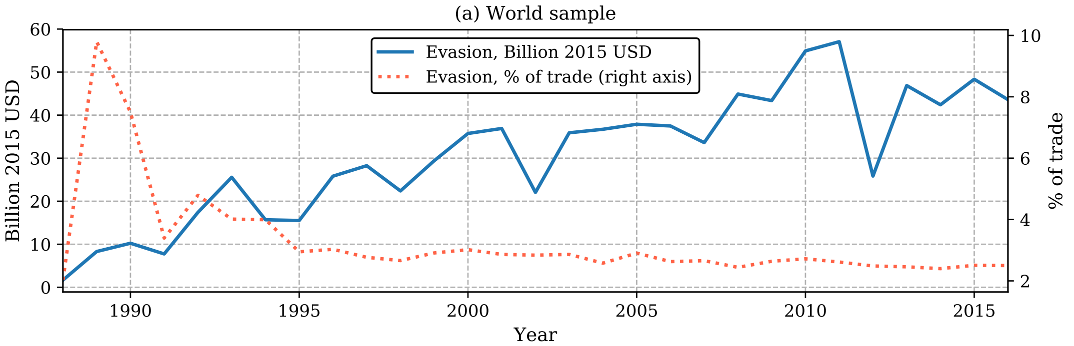
Note: Histograms of (a) tariffs values, for tariffs smaller or equal to 40%; (b) number of changes per product-country pair (iep) triplet; (c) size of tariff changes, dropping changes below and above the 0.1 and 99.9 percentiles, respectively, for the purpose of exposition. Data: world sample.

FIGURE 1.B.16: TARIFF EVASION OVER TIME AND ACROSS INCOME LEVELS

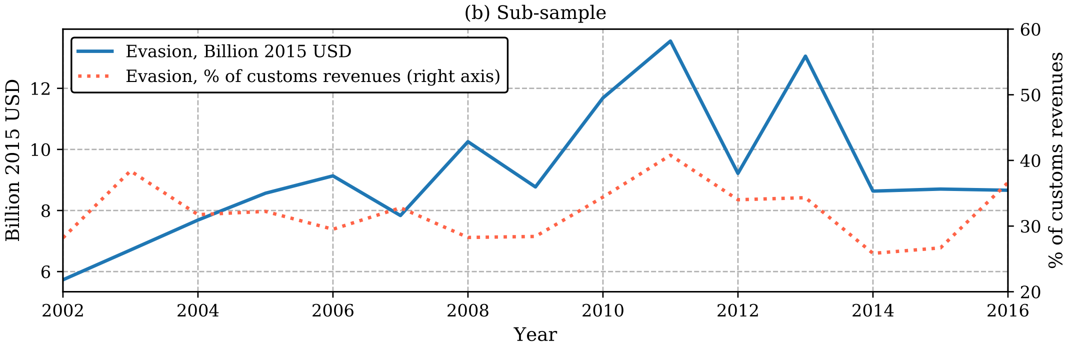


Note: Coefficients on tariffs in regressions based on equation (1.14), where tariff is interacted with categorical variables indicating time periods and income levels. Regression coefficients are from tables 1.A.27 and 1.A.28. The horizontal dashed line represents the baseline estimate of α, the coefficient on tariff in equation (1.14).

FIGURE 1.B.17: AMOUNTS OF TARIFF EVASION OVER TIME



(b) Sub-sample



Note: Estimated amounts of tariff evasion based on equation (1.20). Panel (a) displays total tariff evasion across countries over time (in blue) and the average evasion expressed in percentage of trade subject to tariffs (in orange). Panel (b) displays total evasion for a sub-sample of 19 countries over 15 years, where data on customs revenues is available consistently over time. The orange dashed line represents the average evasion expressed as a percentage of customs revenues across countries, over time. Countries in the sub-sample are: Argentina, Bangladesh, Armenia, Brazil, Costa Rica, El Salvador, Georgia, Guatemala, Jordan, Korea, Malaysia, Mauritius, Moldova, Morocco, Namibia, Nicaragua, Peru, South Africa and Uruguay. Source: world sample and WB's measure *Customs and other import duties* (ID: GC.TAX.IMPT.CN).

Appendix 1.C Additional details on the model

1.C.1 Proof of proposition 1

Consider the following accounting problem, where $C(\cdot)$ is a twice continuously differentiable function.

$$\max_{\{F_r, M_r\}} \quad (F - M) - tM_r - \tau \left[F_r - M_r(1+t)\phi\right] - bC(F_r, M_r, \boldsymbol{\omega}) \tag{1.21}$$

where all the variables and the ranges of parameter values are defined as in the main text. Denote $\boldsymbol{\theta} = (t, \tau)$ the vector containing all tax rates. Taking first order conditions yields the following system that implicitly characterizes the solution $(F_r^\star, M_r^\star)$:

$$\begin{aligned} -\tau - bC_1(F_r^\star, M_r^\star, \boldsymbol{\omega}) &= 0 \\ -t + \tau(1+t)\phi - bC_2(F_r^\star, M_r^\star, \boldsymbol{\omega}) &= 0 \end{aligned}$$

where C_i denote the derivative of $C(F_r, M_r, \boldsymbol{\omega})$ with respect to its ith argument. Define $F_r^\star = f(\boldsymbol{\theta}, \boldsymbol{\omega})$ and $M_r^\star = g(\boldsymbol{\theta}, \boldsymbol{\omega})$, the explicit solutions to (1.21), i.e. such that the system of FOCs is satisfied. Denote the FOCs system as

$$I(\boldsymbol{\theta}, \boldsymbol{\omega}, C_1(F_r^\star, M_r^\star, \boldsymbol{\omega})) = 0, \tag{1.22}$$

$$J(\boldsymbol{\theta}, \boldsymbol{\omega}, C_2(F_r^\star, M_r^\star, \boldsymbol{\omega})) = 0. \tag{1.23}$$

The objects of interest are the derivative of the solutions with respect to elements of $\boldsymbol{\theta}$. Specifically, taking the derivatives of the system of FOCs with respect to element k of $\boldsymbol{\theta}$ can be written in the following form:

$$bH_C \times D_k = A_k$$

where $H_C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$ is the Hessian of the cost function where $C_{12} = C_{21}$, $D_k = \begin{pmatrix} \partial F_r^\star / \partial k \\ \partial M_r^\star / \partial k \end{pmatrix}$ is the vector of partial derivatives of the solutions with respect to parameter k, and A_k is a vector gathering terms of $\partial I(\cdot)/\partial k$ and $\partial J(\cdot)/\partial k$ that do not contain elements of H_C. Solving for D_k yields

$$D_k = b^{-1} H_C^{-1} \times A_k = \frac{1}{b(C_{11}C_{22} - C_{12}C_{21})} \begin{pmatrix} C_{22} & -C_{12} \\ -C_{21} & C_{11} \end{pmatrix} A_k \tag{1.24}$$

Finding derivatives of the solutions boils down to specifying A_k for each parameter k of interest and plugging them in the formula above. For $k = t, \tau$, A_k is as follows:

$$A_t = \begin{pmatrix} 0 \\ -1 + \tau\phi \end{pmatrix} \qquad A_\tau = \begin{pmatrix} -1 \\ (1+t)\phi \end{pmatrix}$$

Once D_t and D_τ are calculated as above, proposition 1 immediately follows.[75] Note that H_C is positive definite and that $C(\cdot)$ is therefore convex under the conditions given in proposition 1. This is because $Tr(H_C) = \lambda_1 + \lambda_2 = C_{11} + C_{22} > 0$ and $\det(H_C) = \lambda_1\lambda_2 > 0$ so it must be that $\lambda_1 > 0, \lambda_2 > 0$ and H_C is positive definite (λs refer to the eigenvalues of H_C).

1.C.2 A generic cost functions that can depend on the tax rates

So far, the cost function was not allowed to depend on the tax rates. However, it is possible that the expected cost of evasion changes together with the tax rate, if for instance tax authorities are more vigilant when taxes are higher, or if the cost depends on the amount of taxed evaded — that themselves directly depend on the tax rates. If that is so, a higher tax rate will not only increase the marginal benefits of evasion, but also the marginal cost, making the signing of partial derivatives for comparative statistics purposes harder.

Given the framework above, the only difference is to allow the cost function to depend on elements of $\boldsymbol{\theta}$. The only term affected in expression (1.24) is A_k, that now becomes:

$$\tilde{A}_k = A_k - b \begin{pmatrix} C_{1k} \\ C_{2k} \end{pmatrix}$$

where $C_{1k} \equiv \partial^2 C/\partial F_r \partial k$ and $C_{2k} \equiv \partial^2 C/\partial M_r \partial k$. Signing the derivatives become harder as more terms are involved and more assumptions need to be imposed. An illustration is provided with how optimal reporting of purchases change as the tariff rate increases, which is given by the following expression:

$$\frac{\partial M_r^\star}{\partial t} = \frac{C_{11}(\phi\tau - 1) + b(C_{21}C_{1t} - C_{11}C_{2t})}{b\det(H)}.$$

The first term in the nominator is negative and the denominator is positive by proposition 1, which, ignoring the second term in the nominator, would result in a negative partial derivative: as tariffs increase, reported purchases decrease. The second term is new and appears as the cost function can now directly depend on the tariff rate t. For the sake of the argument, suppose that $C_{1t} = 0$, i.e. that the change in the cost resulting from misreporting sales — which are not subject to tariffs — is independent of the tariff rate, and that $C_{2t} < 0$, i.e. that the cost of under-reporting purchases increases at t increases — capturing the idea that border authorities scrutinize that good more carefully as the tariff increases.[76] In that case, the second term in the nominator

---

[75] With the exception of showing that $\frac{\partial(F_r^\star - M_r^\star(1+t)\phi)}{\partial\tau} < 0$, which is not obvious. This partial derivative equals $(1\ ,\ -(1+t)\phi)D_\tau$, whose sign equals the sign of the following expression: $-C_{22} - C_{11}((1+t)\phi)^2 - 2C_{12}(1+t)\phi$, which can be shown to be negative. First, rewrite this as $-C_{11}\left[\left((1+t)\phi + \frac{C_{12}}{C_{11}}\right)^2 - \frac{C_{12}^2}{C_{11}^2} + \frac{C_{22}}{C_{11}}\right]$, then notice that $-\frac{C_{12}^2}{C_{11}^2} + \frac{C_{22}}{C_{11}}$ is positive since $\det(H_c) > 0$.

[76] It seems reasonable to assume that $C_{2t} < 0$, $C_{1\tau} < 0$, $C_{2\tau} > 0$, i.e. that the cost function becomes steeper as the tax rate increases over the region where the optimal reported amount is expected to lie. For instance with tariffs, one expects reported purchases to be lower than the true amount, so the cost function should become steeper (more negative) for reported amounts lower that the true value as the

becomes positive and the sign of the derivative depends on which of the two terms is larger. The two countervailing forces are as follows: an increase in t induces firms to under-report purchases more, yet since t increases and customs pay more attention to trade in that good, firms under-report less. Overall, the relative magnitude of these forces determines what the ultimate effect of tariffs on reported purchases will be. Similar analyses can be performed for the effect of CIT on optimal reported amounts. In general, a sufficient condition for the signs of the comparative statics of interest to remain unchanged is that C_{1k} and C_{2k} are small in magnitude.

1.C.3 A model with domestics sales and purchases

Firms generally do not export all their sales nor import all their intermediate inputs. Consider the profit function of a firm that sells and sources both domestically and abroad (denoted with the superscripts D and A, respectively):

$$\pi = (F - M) - tM_r^A - \tau \left[F_r^D + F_r^A - (M_r^A(1+t) + M_r^D)\phi \right], \tag{1.25}$$

where F and M are now the sums of their domestic and foreign components: $F = F^D + F^A$ and $M = M^D + M^A$. The firm can in principle misreport both domestic and trade amounts. What determines the extent of misreporting in each amount is the cost function. Suppose for a moment that the marginal costs to misreport M^D versus M^A, as well as F^D versus F^A are the same. In that case, the model does not have a unique solution since the marginal benefit (in terms of taxes evaded) is the same for both F_r^D and F_r^A. Misreporting in terms of purchases will be a corner solution with all the cheating done on either M_r^D or M_r^A depending on the levels of t, which determines which purchases is best for misreporting.

Allowing misreporting solely through trade Consider the case where misreporting is only possible through trade:

$$\pi = (F - M) - tM_r^A - \tau \left[F^D + F_r^A - (M_r^A(1+t) + M^D)\phi \right], \tag{1.26}$$

with $F = F^D + F^A$ and $M = M^D + M^A$. The marginal benefits, i.e. the derivative of the profit function with respect to reported imports and exports, are identical to the baseline model in the text. Insofar as the cost function remains unchanged, proposition 1 remains unchanged.

The predictions may be altered when the cost function is allowed to depend on the levels of domestic purchases and sales. Specifically, there may be *scale effects* if large volumes of domestic transactions render misreporting trade easier. There is a simple intuition behind this: a firm that sells a large amount domestically may cheat *more* on exports, because this hidden by the scale of domestic sales. This would allow for interesting testable predictions and is left for future work.

---

tariff rate increases.

1.C.4 An illustrative cost function

The firm can evade both tariffs and corporate income taxes by misreporting the values of its sales and purchases. This entails a cost, as detected evasion is punished. Denoting F_r and M_r the reported amounts, the firm faces the following expected cost of evading:

$$C(F_r, M_r) = \frac{b}{2}\left[\left(\frac{F-F_r}{F}\right)^2 F + \left(\frac{M-M_r}{M}\right)^2 M + a\left(\frac{(F-M)-(F_r-M_r)}{F-M}\right)^2 (F-M)\right], \tag{1.27}$$

where $b > 0, a \geq 0$ are parameters capturing the rate at which tax authorities detect evasion. The cost of evading increases when sales or purchases are not reported accurately, and when the reported gross margin differs from the truth. Specifically, the cost is increasing in the percentage deviations of the reported amounts from their true values, and proportional to the true values. This specific cost function accounts for the fact that the firm's accounts must remain consistent: if a firm reports low sales and high costs to lower its taxable revenues, the risk it will be detected increases. This risk is higher if tax authorities can observe and track this gross margins well, captured by a high value of a. The key is that irrespective of the incentives to misreport sales and purchases, they cannot be too far apart. In a sense, the third term of the cost function *binds* reported sales and purchases together.

Given the expected cost of evading, the firm optimally chooses reported sales and purchases to maximize its after-tax profits — which is essentially a compliance optimization problem:

$$\max_{\{F_r, M_r\}} \quad (F-M) - tM_r - \tau \max\{0, [F_r - M_r(1+t)\phi]\} - C(F_r, M_r). \tag{1.28}$$

Under-reporting sales lowers the corporate tax liability, whereas under-reporting imports lowers the tariffs owed yet increases taxable profits simultaneously. The optimal reported sales and purchases are given by:

$$F_r^\star = F + F\left[\frac{aM(\eta-\tau) - \tau(F-M)}{b(F-M+a(F+M))}\right] \tag{1.29}$$

$$M_r^\star = M + M\left[\frac{\eta((1+a)F-M) - aF\tau}{b(F-M+a(F+M))}\right] \tag{1.30}$$

where $\eta \equiv \tau(1+t)\phi - t$ is the marginal effect of reported purchases on profits (ignoring evasion costs) — a higher value of M_r lowers taxes paid through higher costs, yet increases the tariff liability. A positive η implies that over-reporting purchases will increase profits (ignoring evasion costs). Some intuition can be gained from the expressions above. First, each reported amount consists of the true amount plus a markup that can either be positive or negative depending on the tax rates and the parameters of the cost function. Second, as the taxman efficiency at detecting evasion (b) increases to infinity, the optimal reported values tend to the true values. Third, assuming that tax authorities do not monitor the gross margin at all ($a = 0$), reported sales only depends on the CIT rate, b and the true value of sales, and reported purchases only depend on η, b and the true value of purchases. So the dependence of $F_r^\star$ on M and of $M_r^\star$ on F comes from that fact that discrepancies of the reported gross margin from the true margin increases the probability of getting caught via the third term of the cost function. Keeping that in mind, the following comparative statics are proposed.

Proposition 2. *Provided that $F > M > 0$ and $a < (F - M)/(Mt)$:*[77]

1. *Reported sales, $F_r^\star$, depend negatively on the corporate income tax rate;*
2. *Reported purchases, $M_r^\star$, depend negatively on tariffs;*
3. *Reported purchases, $M_r^\star$, depend negatively on the corporate income tax rate if $(1+t) < \dfrac{aF}{\phi((1+a)F - M)}$, and positively if the reverse is true;*
4. *Reported taxable profits, $F_r^\star - M_r^\star(1+t)\phi$, always decrease as a result of an increase in corporate taxes.*
5. *The sensitivity of reported purchases, $M_r^\star$, to tariffs decreases as the CIT rate increases.*

Proposition 2 details the main predictions of the model. Reported purchases decrease when the tariff rate increases, as the firm's incentives to evade tariffs increase. When the CIT rate increases, the firm would like to under-report sales and over-report purchases. However, this would result in a large decrease in the reported gross margin relative to the true gross margin, increasing the expected cost of evasion. Therefore the firm will report sales and purchases in such a fashion that this discrepancy is not excessive. This could lead the firm to lower sales and purchases simultaneously even though lowering purchases in itself results in higher taxable profits.[78] This happens when

$$(1+t) < \frac{aF}{\phi((1+a)F - M)}, \tag{1.31}$$

i.e. when tariffs are low, not all costs are tax-deductible or the attention of tax authorities to discrepancies in the gross margin is high. All three factors imply that over-reporting purchases are relatively unattractive: the firm is better off lowering sales and lowering purchases as well (but by less) to minimize the risk of being caught.[79] When $(1+t) \geq aF/(\phi((1+a)F - M))$, the firm finds it more efficient to inflate costs in order to evade corporate taxes, despite the resulting discrepancy in the reported gross margin and the associated increase in the evasion cost. This is because higher tariffs and more cost deductibility imply a higher marginal benefit of inflating costs.[80]

---

[77] The condition $a < (F - M)/(Mt)$ is sufficient to rule out over-reporting of sales upon an increase in the CIT rate, which could happen in theory for the following reason: CIT rate increases, reported purchases increase, and so do reported sales to avoid a large discrepancy in gross margin, which would increase the expected cost of evasion. The condition $F > M > 0$ ensures positive profits before tax. The predictions of the model should remain unchanged in case of losses, if the losses can be carried forward to the next financial year.

[78] In that case, one can think of the firm reporting in such a fashion as to appear smaller than it actually is.

[79] It can be shown that $\frac{\partial F_r^\star}{\partial \tau} < \frac{\partial M_r^\star}{\partial \tau}$ always. Since the first term is negative, the second term is less negative — or positive when $(1+t) \geq aF/(\phi((1+a)F - M))$.

[80] It is in principle possible that a firm may want to inflate costs *and* inflate purchases (by less) to lower its taxable profits. This is impossible given the condition that $a < (F - M)/(Mt)$. Intuitively, it is easier to imagine a firm lowering its sales and purchases, e.g. by hiding part of its business, rather than creating imaginary sales and purchases for the purpose of evasion.

Last, the magnitude of tariff evasion decreases when the CIT rate increases as a higher corporate income tax rate makes tariff evasion — which results in lower reported purchases and therefore higher taxable profits — less attractive. Stated differently, higher tariffs make condition (1.31) less likely to hold and therefore more likely to over-report purchases upon an increase in the CIT rate.

Proposition 3. *When the tax authorities effectiveness (b) increases:*

1. *Both reported sales and purchases converge to their true value.*
2. *Both reported sales and purchases react less in response to a change in the corporate income tax rate.*
3. *Reported purchases react less in response to a tariff change.*

Proposition 3 details how the reported values change depending on the effectiveness of tax authorities. The magnitude of evasion decreases when tax authorities are more efficient at detecting it.

Appendix 1.D Details on custom valuation

The transaction value is an important object as it determines the amount of duties to be paid by the importer, since most tariffs are *ad valorem*, i.e. expressed in percentage of the transaction value. Members of the WTO must adhere to the ”Agreement on Implementation of Article VII of the General Agreement on Tariffs and Trade 1994”, also known as the Custom Valuation Agreement, CVA henceforth.[81] The agreement specifies that customs valuation must be based on the actual price of the goods imported, in principle as shown on the invoice of the transaction. Importantly, the documentation used to determine the transaction value is provided by the importer. This valuation method prevents custom officers to price imports in a discretionary fashion, or to use external reference prices. If the dutiable value cannot be assessed using the transaction value method, a series of alternatives can be applied, consisting in obtaining a price from identical or similar goods (see here for more details, last accessed 19/05/2020).

This arguably limits the scope for evasion by the importer. Javorcik and Narciso (2017) show that quantity misreporting and outright misclassification are more attractive evasion methods for WTO members relative to price misreporting, as the valuation requires the exporter's invoice on which the price is stipulated. The authors argue that there seems to be more discretion possible for non-WTO countries, in which custom officers may price imports according to external reference prices or their own experience or data. In this case, the importer may possibly bribe customs officers to lower the transaction value and thus the due duties. However, 164 countries are WTO members by 2017 — most of the countries in the sample — and more than 90% of observations involve importers members of the WTO.

---

[81] Prior to 1994 and the creation of the WTO, the 1979 Tokyo Round Valuation Code was used by more than 40 members of the GATT (General Agreement on Tariffs and Trade). The determination of the transaction value was essentially the same as under the 1994 CVA.

Chapter 2

A Geometry of Innovation

2.1 Introduction

Innovation is a key determinant of economic growth. In this paper, we propose a set of measures of the importance of inventions based on their similarity to past and future innovative output. Our thoughts are guided by the simple intuition that inventions that are dissimilar to existing technologies, yet similar to future innovations may have anticipated or created shifts in innovation topics and thus been particularly impactful on the technological landscape.

Our measures require accurate quantification of similarities between inventions. To this end, we represent around 4.6 million patents as numerical vectors in high dimensional space based on their text. Patents of similar textual content are close to each other in vector space. Armed with these vector representations, we analyse the similarity of a patent's content to existing patents and topics (i) at the time when it is filed, and (ii) 10 years later. We test whether patents dissimilar to existing patents/topics when they are filed yet similar to subsequent patents/topics 10 years later are successful innovations, where success is measured in terms of citations, private economic value and growth in output and profitability of the filing firm. We also analyse long-term trends in these scores to understand whether such precursor innovations are harder to produce over time, conditional on research effort. The success of this exercise hinges on how well we can proxy technological differences via differences in the text of patents describing the protected invention.

Data and methodology Our data consists of approximately 4.6 million utility patents granted by the United States Patent and Trademark Office (USPTO) over the years 1975-2017, a subset of which is linked to firms from the Compustat database. After building a document term matrix from their texts — i.e. a word count per document in the form of a matrix — we apply truncated singular value decomposition to it (also often referred to as Latent Semantic Analysis, Deerwester et al., 1990) to obtain vector representations of sizes that are manageable in terms of computing cost. We use these vector representations to calculate three key metrics (scores) for each patent based on its location within the technological space over time. Each score is based on two ingredients: the similarities of the patent to past and future patents. For these comparisons, we focus our attention on patents filed 10 years before and after the filing date of the patent of interest, which form what we call *backward* and *forward spaces*, respectively.

The first score is based on what we name *centroid patents*, which are imaginary *mean* patents computed as the mean of all patent vectors within subsets of the technological space. Specifically, centroid-based scores are computed for each patent by taking a ratio of its distance to the centroid of the forward space over that to the centroid of the backward space, where the spaces either contain patents from all or specific technological fields. A high score implies that a patent was dissimilar to the existing mean patent when filed, yet similar to the subsequent mean patent. In other words, it anticipated a technological shift in the economy or a specific industry.

The second metric — that we call *widening* score — is a refinement of the idea above, where the similarity to centroids is replaced by the average similarity to the patent's closest neighbours. The term *widening* refers to the idea that a patent that is filed in an empty part of space (i.e. that is dissimilar to its neighbours in the backward space) yet becomes central subsequently (i.e. similar to its neighbours in the forward space) has *widened* knowledge in a novel way.

To visualise the spatial geometry of patents, we additionally use a recent t-Distributed Stochastic Neighbour Embedding (t-SNE) variant by Linderman et al. (2019) from RNA sequencing visualisation in genetics (see also Maaten and Hinton, 2008, for the first paper on this method). t-SNE is a tool that allows the visualisation of data patterns present in high-dimensional space in low-dimensional spaces of two or three dimensions. Applied to our patent representations, it allows us to see clusters and neighbourhoods that relate to different technological fields and it visually guides our analysis.

Findings We find that patents which anticipated economy-wide and field-specific shifts are cited more, and have higher private economic value. At the firm level, firms which were granted patents that anticipated shifts in their field tend to grow faster and to be more profitable, although our analysis does not allow us to claim any causal links between the specific patent we identify and firm outcomes due to the existence of differential pre-trends. However, we show that it is possible to identify innovative firms, which also seem to make higher profits, purely based on the language content of patents. Due to the persistence of R&D quality in firms, these firms also have been performing better than their peers some years before filing the patent of interest.

Patents which we identify as *widening* innovations also tend to be cited more and have a higher private value. Similarly to patents anticipating more systemic shifts, firms that patent these inventions tend to have higher growth in profits and output, as well as capital. Again, firms that we identify as innovative based on their patents' spatial properties also grew faster than other firms a few years prior to filing the identified patents. It it reassuring to see that the inventions our methods pick among close to 3 million patents are the intellectual output of fast-growing, successful firms.

In a subsequent discussion, we argue that methods like ours largely pick up the evolution of technology over the past 30 years and in particular the information technology (IT) revolution, which we know has significantly shaped the economy. The IT revolution is identified by all scores: it resulted in economy-wide and field-specific shifts in the content of innovation, and early IT patents in the 1990s appeared in a completely empty part of our representation of the technological space. We find that IT is responsible for a considerable share of the results we present in this paper, yet the results qualitatively survive in a sub-sample dropping the most IT-intensive fields of technology.

Last, trends over the sample period reveal that in fields with high average widening scores, which we denote *higher-innovation fields*, the number of patents filed increased consistently over the whole sample period, and so did R&D expenditures. However, patents per dollar of R&D spending plateaued after 1995. Furthermore, patents in these fields become increasingly similar to their closest neighbours over time, and while the average widening score increases up to 1995, it sharply decreases afterwards. We tentatively interpret these patterns as suggestive that ideas may have gotten harder to get over time in these fields and provide a discussion of the limitations of this interpretation.

Relation to the literature This paper relates to several strands of the literature. First, it contributes to the extensive literature that applies NLP methods to patents to characterize technological progress. Balsmeier et al. (2018) and Packalen and Bhattacharya (2015) identify novel inventions based on the first appearance on words in the patent corpus. Bowen, Frésard and Hoberg (2018) construct a measure of ex-ante technological disruptive potential of patents based on their use of new or fast-growing words across contemporaneous patent applications. Arts et al. (2018) develop a measure of text-based patent similarity and ask technology exports from different fields to validate it. The closest paper to the present work is a very recent working paper by Kelly, Papanikolaou, Seru and Taddy (2018), who develop a conceptually identical method to ours to identify significant inventions, based on their similarity to previous and subsequent patents.[1] Our paper's widening score is most closely related to their score. However, we also develop innovation scores based on mean patents and provide illustrations of the technological space via novel implementations of dimension reduction techniques that preserve data patterns from high-dimensional space in low-dimensional space (Linderman et al., 2019). We also provide tentative evidence that ideas may have gotten harder to get over time, in line with findings from Bloom et al. (2020) — among others.

This work also relates to the literature on innovation and economic growth. First, it contributes to attempts at evaluating the economic importance of inventions. This is a difficult task only imperfectly measured by citations, which tend to reflect scientific importance — which is however related to economic value (Hall, Jaffe and Trajtenberg, 2005; Nicholas, 2008). Kogan, Papanikolaou, Seru and Stoffman (2017) provide a measure of private economic value based on the stock market response to news about patent publications.[2] Our measure of the importance of patents is correlated to both citations and private economic values, but provides additional insights as to what features of inventions create value. Our concept of *widening* inventions relates to the notion that an innovation can deepen the specialized knowledge of a given narrowly defined technological field, or widen the scope of knowledge by creating a yet

---

[1] We refer to this study throughout the paper. When specific references to tables and equations are given, they refer to the first NBER working paper version dated November 2018, unless otherwise specified.

[2] This measure is extended by Kline et al. (2019) to a larger sample of US firms.

non-existent technology. It also relates to endogenous growth models in the tradition of Klette and Kortum (2004), where innovations can either replace existing products or create new ones. Our measure of *widening* patents tentatively provides a tangible quantification of the concepts underlying these theories.[3]

Structure of the paper The paper is structured as follows. Section 2.2 describes the data. Details of the methodology are given in section 2.3. Section 2.4 presents the results. A discussion of the importance of the IT revolution for our results, and of several potential measurement issues relating to our methods is provided in section 2.5. Section 2.6 concludes. Additional tables and figures, respectively referenced with prefixes 2.A and 2.B, can be found in appendices of the same name.

2.2 Data

This section contains details on the data used in this project. Data on patents is from the United States Patent and Trademark Office (USPTO). Firm-level data is from the CRSP/Compustat Merged Database, and data on the economic importance of patents is from Kogan et al. (2017) and (Stoffman et al., 2019).

2.2.1 Patents

The USPTO provides data on all patents that have been granted in the US, both by US and foreign entities. We restrict the sample to utility patents granted by the USPTO over the years 1976-2017, for which detailed information in machine-readable format is available. There are some cases where two or more patents have the exact same abstract. This happens when an original patent is followed by continuing applications that either change claims or focus on a subset of claims from the original application.[4] Since these applications refer to the same invention, we only keep the patent that was first granted.

For each patent, the following information is available: the full texts of the abstract, the description of the invention and the claims; the filing and grant dates; the International Patent Classification (IPC), which indicates to which areas of technology

---

[3] Other work measuring the technological breadth and depth of sectors, firms or patents include Katila and Ahuja (2002); Ozman (2007); Moorthy and Polley (2010); Lodh and Battaggion (2014), who base their measures on the diversity and intensity with which technological sub-fields are related to each other. Our measure is purely based on the text of patents.

[4] See section 201 of the USPTO manual, for exact definitions (last accessed 19/05/2020).

TABLE 2.1: DESCRIPTIVE STATISTICS OF PATENT CITATIONS

	N	Mean	Sd	Min	p25	Median	p75	p90	Max
5-year citations	4,633,305	3	7.1	0	0	1	3	7	1431
10-year citations	4,633,305	6.4	16	0	0	2	6	15	2801
15-year citations	4,633,305	8.9	24	0	0	3	8	21	2801
Citations as of 2017	4,633,305	12	32	0	0	3	11	28	3997

Note: descriptive statistics of the number of citations of patents at different horizons after their grant dates.

a patent pertains; information on the filing and beneficiary entities (which can differ); and data on forward citations by subsequent patents.[5] The final sample consists of about 4.6 million patents granted over the years 1976-2017. Out of these, a subset of around 1.67 million patents is matched to firms from Compustat, using the matches publicly provided by Kogan et al. (2017) and Stoffman et al. (2019).[6] These studies also provide an estimate of the economic value of most of these patents.

The final sample of patents for which a score can be computed consists of around 2.8 million patents over the years 1985-2008 — as the score can only be calculated for patents with 10 years of data before and after their filing date, see section 2.3 for details on the methodology. Out of these 2.8 million patents, around 1.15 million can be matched to Compustat firms. Figure 2.B.1 plots the number of patents filed and granted in the sample over time, and the coverage in terms of scores and linkage to firms. In panel (a), note that the number of filed patents decreases dramatically in the later years. This is because the data only contains patents that have been granted and most patents filed in recent years have not been granted yet. On average, a patent is granted 2.5 years after it was filed. In the whole sample, patents are cited 6.4 times on average in the 10 years following their publication. The distributions of forward citations is skewed to the left: 50% of patents get 2 or less citations over 10 years — descriptive statistics of the number of citations at different time horizons can be found in table 2.1.

2.2.2 Firms

Firm-level data comes from the Compustat database. It provides financial and accounting data for listed firms in the US. Following Kogan et al. (2017), we restrict the

---

[5] Throughout the paper, we use the following terminology to indicate the level of aggregation of the IPC we refer to: IPC 1 for the 1-digit level, i.e. sections and IPC 3 for the 3-digit level, i.e. classes.

[6] The data from Kogan et al. (2017) can be found here. Last accessed on 03/05/2020. The data from Stoffman et al. (2019) can be found here. Last accessed on 03/05/2020. We use the private value and the patent-firm matches from Stoffman et al. (2019) in this version of the paper since the coverage of the data is more extensive. The results remain largely unchanged irrespective of the data source we use.

TABLE 2.2: DESCRIPTIVE STATISTICS OF THE FIRM SAMPLE

	N	Mean	Sd	Min	p25	Median	p75	p90	Max
# patents granted	119,836	8.8	78	0	0	0	1	5	4254
# patents filed	119,836	10	98	0	0	0	1	6	8934
1-y gr. rate: capital	107,436	.13	.46	-12	.024	.097	.22	.45	10
1-y gr. rate: profits	98,192	.072	.5	-9.7	-.083	.061	.22	.48	7.8
1-y gr. rate: output	106,319	.082	.51	-10	-.06	.056	.2	.44	9.4
1-y gr. rate: employment	102,799	.05	.41	-8.8	-.062	.029	.15	.35	8.6

Note: # patents granted and filed: number of patents granted and filed for a firm in a given year; 1-y gr. rates: 1-year growth rates of firm outcomes as defined in the text.

sample to observations where values of SIC (*Standard industrial classification of economic activities*) codes and book assets are not missing. We omit industries that never patent, as well firms from the financial and utilities sectors (SIC codes 6000 to 6799 and 1900 to 1949, respectively). The final sample consists of about 120,000 firm-year observations over the years 1985-2008. Around 26% of firm-year observations filed at least one patent over the sample period. The variables of interest at the firm level are profits — defined as sales minus costs of goods sold — output — defined as sales plus inventories — employment and capital (property, plants and equipment). All variables are expressed in real terms — profits and output are deflated using the CPI, whereas capital is deflated using the equipment implicit price deflator from the National Income and Product Accounts (NIPA). All firm-level variables are winsorized at the 1% percent level yearly.

Table 2.2 provides summary statistics for the main variables. The distributions of number of patents filed and granted are very skewed to the left: while most firms do not file a patent, few firms innovate massively, filing in excess of 4000 patents in a given year.[7] The big innovators — in terms of patent filed — are large firms like IMB, Microsoft or Sony, to name a few. Firms in this sample tend to be large (they are all listed companies) and also tend to grow fast. For instance, the average yearly growth rate of output is 8.2%.

2.3 Methodology

The following sections introduce the methodology of this paper. In section 2.3.1 we describe how we build patent representations, illustrate their quality and offer graphical visualisations thereof. We proceed in section 2.3.2 with explanations of the computations of the scores we derive for each patent and on which this paper relies, accom-

---

[7] An exceptional outlier is IMB in 2008, which filed 8934 patents.

panied with illustrative examples. Section 2.3.3 introduces the regression frameworks for the analysis of patent- and firm-level outcomes.

2.3.1 Patent representations

The ultimate aim of this section is to represent patents in the form of numerical vectors based on their textual content. The main purpose of our patent representations is to measure how similar patents are to each other. It may be useful to think of each patent as a dot in a plane, in which case a numerical representation of a patent is the coordinates of that dot. If the texts of two patents are similar to each other, the dots representing them will be close to each other on the plane.

2.3.1.1 Derivation

This section explains how we obtain numerical representations of patent texts (documents) based on the words (terms) they contain. Throughout this section, matrix rows and columns are indexed by $i = 1, \ldots, m$ and $j = 1, \ldots, n$, respectively. We denote row i of matrix B as b_{i*} and column j as b_{*j}

Document term matrix The document term matrix (dtm) stores the number of occurrences of each word (term) for each patent (document) in matrix form. Suppose the entire corpus consists of four patents. Patent one's full text reads "Invention!", patent two reads "A good invention.", patent three "A better invention.", and patent four "A last good invention." When analysing the USTPO database, we concatenate the texts of the abstract, brief description, and claims. After deleting punctuation, numbers, and converting everything to lower case letters, the resulting document term matrix — denoted $\tilde{A}$ — would have the following form, storing word counts for $m = 4$ documents and $n = 5$ terms:

$$\underset{4\times 5}{\tilde{A}}^{toy} = \begin{array}{ccccc} a & \mathit{invention} & \mathit{better} & \mathit{good} & \mathit{last} \\ \hline 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \end{array}$$

Each row vector is a 5-dimensional patent representation. As already visible in this example, dtms are usually sparse, i.e. contain large numbers of zero elements. In our application, we create vector representations for $m = 4,633,363$ unique patents from 1976 to 2017. To reduce noise and irrelevant information, we delete words which appear in less than 5 documents. Furthermore, we also delete all words which are contained in 15% or more of the documents.[8] Without this, each row vector/document would seem similar as they would share very common words ('the', 'invention', etc.) which have little signal for the technical content of a patent.[9] Although both steps substantially decrease the number of columns, we are still left with $n = 765,149$ columns/words in our final dtm, which is therefore of dimension $4,633,363 \times 765,149$. To make documents comparable despite being of different lengths, we divide each row by its total word count to obtain the shares each word represent in the document and denote the resulting matrix A:

$$A = \text{diag}^{-1}(r_1, \ldots, r_m)\tilde{A}$$

where r_i is the sum of all elements of row i of $\tilde{A}$: $r_i = \sum_{j=1}^{n} \tilde{a}_{ij}$. Unlike related work such as Kelly et al. (2018), we do not employ additional inverse document frequency (idf) steps — which weigh words by the inverse of their relative frequency of appearance, i.e. giving more importance to rarer words — but rather delete common words across all years as discussed above. Kelly et al. (2018) also introduce a concept they call "backward-IDF" which makes the idf weights solely depend on previously filed patents (see Kelly et al., 2018, for full details). In contrast, we build the patent dtm using all years of the dataset and subsequently select the rows linked to the years we study. In our example above, normalising by word counts implies the following dtm:

$$\underset{4\times 5}{A^{toy}} = \begin{array}{ccccc} a & invention & better & good & last \\ \hline 0 & 1 & 0 & 0 & 0 \\ 1/3 & 1/3 & 0 & 1/3 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 \\ 1/4 & 1/4 & 0 & 1/4 & 1/4 \end{array}$$

---

[8] We also delete stop-words such as 'the', 'a', etc., however, these words are likely contained in 15% or more of the documents and therefore would already have been deleted.

[9] Note that we abstract from this step in the exemplary dtm discussed here.

Centroids This matrix now also allows the illustration of a concept which we use to direct our thinking in this paper. It is the mean patent vector of (potentially a subset of) the dtm or its counter parts in reduced dimensions (see next paragraph). We call this imaginary mean patent *centroid*, which for any matrix B is the vector of mean values of its columns:

$$centroid_B = (m^{-1}\sum_{i=1}^{m} b_{i1},\ \ldots,\ m^{-1}\sum_{i=1}^{m} b_{in})$$

In the above example the centroid patent vector would be $centroid_{A^{toy}} = (11/48, 23/48, 1/12, 7/48, 1/16)$. Different times and different industries (i.e. subsets of rows) have different centroids, which are essentially the average frequency of occurrence of a term across documents.

Dimension reduction with truncated SVD In the next step, we reduce the dimensionality (columns) of the dtm A. On the one hand this yields more manageable dimensions for computing millions of similarities between vectors, on the other hand it allows parts of the graphical analysis of section 2.4 which is based on low dimensional representations. Following the Natural Language Processing literature, we use truncated singular value decomposition (tSVD) on A to obtain an approximation of it which we store in matrix Z.[10] For a discussion of the technique, see for example Strang (2016). If we had subtracted column means in A before applying the SVD, the columns in Z would be the often-used Principal Components. The truncated version of the SVD computes only those vectors associated with the largest N singular values of A. This makes running this method feasible when considering very large matrices such as our true dtm which is of dimension $4,633,363 \times 765,149$, i.e. has trillions of cells. We choose the first 300 components from the tSVD to represent our documents in reduced dimensions. The matrix Z is therefore of dimension $4,633,363 \times 300$. Each component/column is a linear combination of words. Each row is a depiction of a patent's textual content in a common space. Vectors of patents with similar content have a similar angle. Running such a dimension reduction on a dtm and using it for analysis is often called Latent Semantic Analysis (Deerwester et al., 1990). In fact, a

---

[10] SVD is a factorization of a matrix B which takes the form $U\Sigma V^T$ where the diagonal entries of Σ — which is a diagonal matrix — are the singular values of B, and U and V are the corresponding left- and right-singular vectors of B, respectively. For a given matrix B, Σ is unique when the decomposition is such that the singular values are in descending order. Truncated SDV yields an approximation of B of the form $U_k\Sigma_k V_k^T$ where only the largest k singular values and associated singular vectors are used.

system using it for information retrieval was actually patented in the USTPO under patent number *4839853* from 1988 but it is now expired![11] To continue with our example, we run tSVD on our exemplary 4×5 dtm A^{toy} with three components. It yields:

$$\underset{4\times 3}{Z^{toy}} = \begin{array}{ccc} z^{toy}_{*1} & z^{toy}_{*2} & z^{toy}_{*3} \\ \hline 0.9377 & 0.3453 & -0.0387 \\ 0.4621 & -0.3002 & -0.1103 \\ 0.4390 & -0.1909 & 0.3223 \\ 0.3626 & -0.2791 & -0.1496 \end{array}$$

Dimension reduction with t-SNE We use the vectors obtained from the tSVD for computations. These computations are faster with 300-dimensional vectors than with their 765,149-dimensional counterparts in the dtm. Yet, to visualise the patents in space, we need representations in two or three dimensions. Taking only the first 2 or 3 columns from the SVD would not preserve enough meaningful geometric information. Could we still have a glimpse into the structures in high dimensional space? The method of t-distributed stochastic neighbour embedding (t-SNE) by Maaten and Hinton (2008) achieves exactly this. It preserves cluster and structures from high dimensional space and separates clusters of vectors in low dimensional space. This makes groups and patterns of observations visible. Intuitively, it depicts objects that are similar (dissimilar) in high-dimensional spaces as points that are close (far) from each other in low-dimensional space. We employ a very recent and faster variant of t-SNE developed by Linderman et al. (2019) that the authors used for visualising single-cell RNA sequences in genetics. This fast interpolation-based t-SNE (FIt-SNE) creates a good approximation and is feasible even for matrices of large dimensions like our patent representation matrix containing around 4.6 million patents/rows. We use their method to further reduce the dimensions of the patent representations contained in Z down to only two columns:

$$\underset{4{,}633{,}363\times 300}{Z} \rightarrow \underset{4{,}633{,}363\times 2}{Z^{\text{t-SNE}}}$$

Again, each row in this new matrix is a patent that can now be represented by only two values (for visualisation only). We can therefore plot its position in a plane. In terms of parametrisation, we use 20,000 iterations and a learning rate of 386,113

---

[11] See `https://en.wikipedia.org/wiki/Latent_semantic_analysis`

which is based on the amount of rows in our matrix.[12]

Computing similarities between patents The main purpose of our patent representations is to compute similarities between patents. Patents of similar content should have similar numerical vectors. Instead of the row vectors contained in A — which are too large for computation purposes — we use those in Z to compute these similarities.[13] In line with standard practice in the NLP literature, we use the cosine similarity to measure the similarity between two vectors — i.e. the angle between them — as opposed to the Euclidean distance which is undesirably sensitive to vector lengths.[14] Cosine similarity between vectors x and y is given by:

$$\text{cosim}(x, y) = \frac{\sum_{k=1}^{n} x_k y_k}{\sqrt{\sum_{k=1}^{n} x_k^2}\sqrt{\sum_{k=1}^{n} y_k^2}} \in [-1, 1] \tag{2.1}$$

Going back to our example, patent number 2 "A good invention" has the following cosine similarities with the other patents:

$$\text{cosim}(z_{2*}^{toy}, z_{i*}^{toy}) = (0.5941,\ 1,\ 0.69251,\ 0.9900).$$

The patent's similarity with itself is 1 (the Cosine of a 0-degree angle is 1) and the most similar patent is indeed patent 4 "A good last invention." which only differ by one word.

2.3.1.2 Illustrations

Illustrating match quality Once we have computed Z, each row vector is the representation of a patent. As discussed above, picking one row, computing its similarity with each other row, and then sorting the array allows to find the most similar patents. Table 2.3 illustrates the most similar patent found for a few examples. Sometimes, the

---

[12] We are grateful to the author George Linderman for his explanations and the recommendation of the paper Kobak and Berens (2019) which, among others, discusses hyper-parameter settings for t-SNE. We use their recommendation for the learning rate ($m/12$) and leave all other hyper-parameters at FIt-SNE's default with the exception of the number of iterations which we set to $20,000$.

[13] We cannot use row vectors contained in $Z^{\text{t-SNE}}$. These vectors are only used for graphical illustrations as, in particular, the distances between vectors in 2 dimensions (i.e. points) are distorted to allow for clusters of similar vectors from higher dimensions to be depicted graphically. See figure 2.1 for a practical example.

[14] When thinking about similar patents, it might seem most natural to think of Euclidean distance: vectors which are close in space should have similar content. This works if we standardise vectors by their L2 norm, forcing them to have unit length and thus to lie on an N-sphere. The reason is that similar documents, i.e. rows in Z, might otherwise still have different lengths. However, the angle of vectors that point into the same direction is zero, even if these vectors have different lengths, hence the appeal of Cosine similarity.

closest match is a patent which seems to be a small modification of the original (see e.g. example 3 in the table). Yet, we also show cases where our tools returns a most similar patent which is a less close match. Match quality seems very good given that such a method only has information on the set of words that a document contains. In fact, Latent Semantic Analysis has also been proposed for patent examination support (Elman, 2007). Example 4 is the only somewhat odd match. In fact, both technologies are much more similar than visible from the abstracts and can e.g. be used for earthquake detection.[15] This is why we concatenate the texts of abstract, brief description, and claims instead of solely relying on the abstract. Furthermore, methods like ours would be biased by common headers or names contained in the patents, hence the deletion of frequently appearing words (those that are in at least 15% of documents in our case). Yet, matches would still be biased by headers or other common structures contained in small subsets of patents and which could be unrelated to the inventions themselves, e.g. if one firm uses a specific and unique writing style for all of its patents.

A further illustration of the match quality is provided in figure 2.B.2. On average, the 100 closest neighbours of any given patent are more likely to come from the same technological field (as measured by the IPC 1, i.e. IPC sections) than any other field. This is reassuring as patent contents are likely to be field-specific.

Illustrating t-SNE The 300 components of Z obtained via tSVD appear to be good representations of documents which allow us to calculate similarities between patents. Yet, two of those alone unfortunately do not contain enough *visual* information to discern any interesting structures from high dimensional space in just two dimensions. Figure 2.1 illustrates this. It contains all patents (each being a point on the plane) from 1976 to 1985 represented by their first two vectors from the tSVD — in panel (A) — and from the t-SNE — in panel (B). The difference is very stark. t-SNE, which was built to separate cluster structures from higher dimensions into lower dimensional representations, allows groups of patents of similar content to emerge graphically. We will later show that patents within clusters come from similar technological fields.

---

[15] This is visible from the full texts of the patents, which can be accessed online here (last accessed 19/05/2020).

TABLE 2.3: EXEMPLARY PATENT MATCHES

Reference patent	Beginning of abstract	Closest match	Beginning of abstract
9475668 (2013)	A modular element for a creel includes a structure having at least one support for supporting a package or bobbin of yarn; the structure being modularly couplable with other similar structures to allow the feeding of multiple yarns to a textile machine; ...	4753064 (1986)	The spinning or twisting machine comprises a plurality of drafting rolls, a plurality of associated spindles and positioned between them a plurality of yarn breaking devices pivotable between an upright spinning machine operating position into a yarn breaking position. ...
4900994 (1988)	An automatic window glass elevating apparatus for moving a motor normally or reversely by drive control by a switch actuation to move a window glass in a closing or opening direction which has a resistor connected in series with the motor, ...	4746845 (1986)	An automatic window regulator for automobiles includes a drive motor for lifting and lowering the window glass, and a detector for detecting when there is a foreign object jammed between the window glass as it is lifted or lowered and a window frame. ...
4640147 (1987)	A gear assembly comprising two gears and a spring in the form of a C-shaped clip interconnecting the two gears. Two pins are provided on one of the two gears and the spring has two holes, one in each end thereof, whereby the spring can be carried by one gear in a pretressed state by means of the pin-and-hole connection. ...	4745823 (1988)	A gear assembly comprising two gears and a spring in the form of a C-shaped clip for applying a resilient force between the gears. Two pins are provided on the respective side surfaces of the gears to receive the spring which has concave end surfaces to be received by the pins. ...
8919201 (2012)	An acceleration measuring apparatus that can easily detect acceleration with high accuracy is provided. In the apparatus, positional displacement of a swingable pendulum member is detected, feedback control is performed to maintain the pendulum member in a stationary state using an actuator, ...	8677828 (2012)	Provided is a device capable of easily and accurately detecting a vibration period when, for example, an earthquake occurs. When a quartz-crystal plate bends upon application of a force, capacitance between a movable electrode provided at its tip portion and a fixed electrode provided on a vessel to face the movable electrode changes, ...

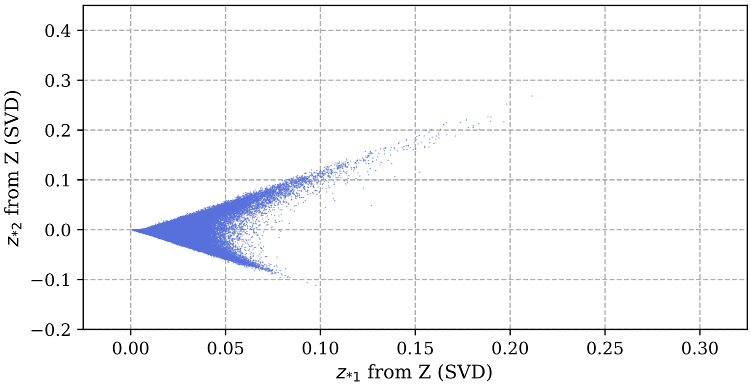
Note: first sentences of the abstracts of 4 patents (left column) and that of their closest match (right column). The 7-digit numbers to the left of the texts correspond to the patent number, and the file year is in parenthesis underneath.

2.3.2 Computing scores

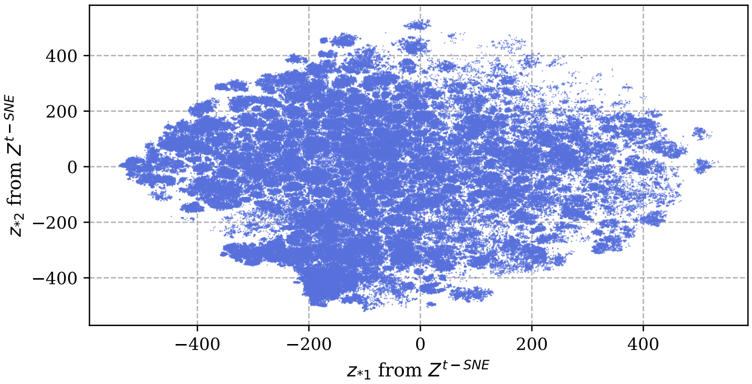
The empirical analysis of this paper employs a range of metrics which proxy the inventiveness of a patent. This section describes the calculations of these innovation scores which we compute for each patent, continuing the illustrative examples from

Figure 2.1: Difference between tSVD and t-SNE

(A) The two first components obtained from the SVD



(B) The two first components obtained from the t-SNE



Note: each dot represents a patent filed from 1976 to 1985. Panel (A) plots the two first components (columns) from Z obtained via tSVD. Panel (B) plots the two first components (columns) from $Z^{\text{t-SNE}}$ obtained via t-SNE.

the previous section.

Backward and forward spaces Each score is based on two fundamental ingredients. A comparison of the patent of interest to patents in the past, and a comparison to patents in the future. For this reason, we introduce the notions of backward and forward spaces, which are sets of patents filed within a number of years of the patent of interest. Both include patents from the current year, i.e. all patents filed in the same year as the patent of interest. Both our backward and forward time intervals consist of 10 years.[16] Going back to our example, if the 4 patents in Z^{toy} are from 1976, 1989,

---

[16] If a patent is from 1995, then its backward interval would be $\{1986, 1987, \ldots, 1995\}$ and its forward interval would be $\{1995, 1996, \ldots, 2004\}$. The intersection of the two sets is $\{1995\}$.

1995, and 2004:

$$\underset{4\times 3}{Z}^{toy} = \begin{array}{cccc} z_{*1}^{toy} & z_{*2}^{toy} & z_{*3}^{toy} & \\ \hline 0.9377 & 0.3453 & -0.0387 & (1976) \\ 0.4621 & -0.3002 & -0.1103 & (1989) \\ 0.4390 & -0.1909 & 0.3223 & (1995) \\ 0.3626 & -0.2791 & -0.1496 & (2004) \end{array}$$

Then for patent number three from 1995 we have the following backward and forward space, which we denote Z^b and Z^f:

$$\underset{2\times 3}{Z}^{b,1995,toy} = \begin{array}{cccc} z_{*1}^{toy} & z_{*2}^{toy} & z_{*3}^{toy} & \\ \hline 0.4621 & -0.3002 & -0.1103 & (1989) \\ 0.4390 & -0.1909 & 0.3223 & (1995) \end{array}$$

$$\underset{2\times 3}{Z}^{f,1995,toy} = \begin{array}{cccc} z_{*1}^{toy} & z_{*2}^{toy} & z_{*3}^{toy} & \\ \hline 0.4390 & -0.1909 & 0.3223 & (1995) \\ 0.3626 & -0.2791 & -0.1496 & (2004) \end{array}$$

In our data, these two spaces can easily contain hundreds of thousands of patents. Taking a constant 10 year interval around a given year makes computations comparable across patents, yet prevents us from computing scores for any patent filed before 1985 and after 2008. In some applications, backward and forward spaces are limited to patents filed within the same technological fields as the patent of interest. In the next sections, we detail the calculations of each score.

2.3.2.1 Centroid-based scores

Since centroids are mean vectors, observing the evolution of centroids of subsets of Z can give insightful information on the importance of topics over time and across technological fields. Centroids are therefore informative in themselves. Our centroid-based scores capture how a patent differs from topics that were important in the past and *anticipates* topics that become important in the future. Centroid-based scores therefore rely on the comparisons of a given patent vector to the centroids of the backward and forward spaces. Continuing with the example, the backward and forward centroids are given by:

$$centroid_{Z^{b,1995,toy}} = (0.4506, -0.2456, 0.1060)$$

$$controid_{Z^{f,1995,toy}} = (0.4008, -0.2350, 0.0864)$$

Following Kelly et al. (2018), we name similarities to patents in the past and future backward and forward similarities, respectively. If a patent is dissimilar to the backward centroid and similar to the forward centroid, we conjecture it is innovative and we would like it to have a high score. Like Kelly et al. (2018), we use the fraction of two similarities as the final score, albeit our scores in this paragraph are based on centroids and not on pairwise comparisons. We denote the score of patent i as Score_i and thus for patent 3 in our toy example:

$$\text{Score}_3^{\text{centroid-based}} = \frac{\text{cosim}(z_{3*}^{toy}, controid_{Z^{f,1995,toy}})}{\text{cosim}(z_{3*}^{toy}, controid_{Z^{b,1995,toy}})} = \frac{0.9117}{0.9221} = 0.9887$$

The score is less than 1, meaning that patent 3 is more similar to the backward space centroid than to the forward space centroid. In the examples above, the backward and forward spaces contain all patents irrespective of their technological field. We name scores relying on all patents *macro scores* and the centroids used in their construction *macro centroids*, as they span all innovations. In contrast, we are also interested in refining the analysis and study dynamics *within* technological fields. In this case, we compute *IPC centroids* and *IPC scores* at the IPC level, i.e. only including patents from the same IPC code as the patent of interest in the backward and forward spaces.

2.3.2.2 Widening scores

The widening score does not rely on centroids, but on the distance of a patent to its closest neighbours in the forward and the backward spaces. *Widening innovations* are patents whose textual content is different from existing patents at the time when they are filed and similar to patents subsequently filed. The term widening reflects the idea that these patents extend knowledge in a novel way and spur innovation. The steps to construct this score are similar to those in building the centroid-based scores. First, we compute the cosine similarity between the patent vector of interest and all patents in the backward and forward spaces (including itself). The similarities are stored in vectors BS_i and FS_i, which stand for backward and forward similarities:

$$
\begin{aligned}
BS_i &= (\text{cosim}(z^b_{k*}, z^b_{1*}),\ \ldots,\ \text{cosim}(z^b_{k*}, z^b_{m*})) \\
FS_i &= (\text{cosim}(z^f_{l*}, z^f_{1*}),\ \ldots,\ \text{cosim}(z^f_{l*}, z^f_{m*}))
\end{aligned}
$$

where the indices k and l correspond to the rows patent i occupies in Z^b and Z^f, respectively. In our toy example:

$$BS_3 = (0.6925, 1) \quad \text{and} \quad FS_3 = (1, 0.5912)$$

Next, the *backward-* and *forward neighbourhood similarities* are defined as the mean distance of the patent of interest to the 100 closest patents in the respective spaces (excluding itself). For this purpose, we store the 100 largest values of BS_i and FS_i excluding the similarities with itself in sets BNS_i^{100} and FNS_i^{100}. Our toy example reaches its limit as the backward and forward spaces only contain one patent each — other than the patent of interest — so in this case:

$$BNS_3^{100} = \{0.6925\} \quad \text{and} \quad FNS_3^{100} = \{0.5912\}$$

Again the final score is built such that a patent that was dissimilar to patents in the past and similar to patents in the future has a high score. We compute the final widening score as:

$$\text{Score}_i^{\text{widening}} = \frac{\overline{FNS}_i^{100}}{\overline{BNS}_i^{100}}$$

where $\overline{FNS}_i^{100}$ and $\overline{BNS}_i^{100}$ are the averages of the sets. In our toy example, the final score is given by $\text{Score}_3^{\text{widening}} = 0.5912/0.6925 = 0.8537$. Again, a score less than one indicates that the patent is closer on average to its neighbours in the past than to its neighbours in the future. Contrary to centroid-based scores, the widening score compares the patent vector of interest to all patents contained in the backward and in the forward spaces. The centroid-based scores rely on the mean patents to describe the backward and forward spaces and involve much less pairwise comparisons, which makes them computationally cheaper. Yet, our widening score should have a better chance at finding which parts of space that are still relatively empty as the patent is compared to all other individual patents in these spaces. Our widening score is almost identical to that developed by Kelly et al. (2018), except that they take the sum of similarities to all patents in the spaces. The next section introduces the empirical strategy we use to link measures of patents' success to their scores.

2.3.3 Regression frameworks

This section details the regressions that are estimated at the patent and firm levels. The regressions will be run with both centroid-based and widening scores. In this section, we therefore us a generic term *score* in the regressions and the descriptions for exposition purposes.

2.3.3.1 Citations

The aim is to check whether our scores are associated with forward citations. Strictly speaking, the exercise is not a prediction exercise as information available only several years after the filing date is used to compute the scores. We estimate regressions of the following form:

$$\text{citations}_{pjf,t+h} = \beta \text{ score standardized}_{pjf,t} + \mathbf{FE} + \epsilon_{pjf,t}, \tag{2.2}$$

where the dependent variable is the number times a patent p from IPC code j owned by firm f — when this information is available — is cited by other patents h years after its grant year t, where $h \in \{5, 10\}$. The variable *score standardized* is the score of the patent, as defined above, standardized to unit standard deviation to make the interpretation of regression coefficients easier. **FE** is a set of fixed effects that varies depending on the exact specification. Combinations of filing year, IPC codes at different levels and firm fixed effects are included. Standard errors are clustered at the filing year level. An alternative specification in logs will also be estimated for robustness, as follows:

$$\log(1\text{+citations}_{pjf,t+h}) = \tilde{\beta} \ \log(\text{score}_{pjf,t}) + \mathbf{FE} + \epsilon_{pjf,t}, \tag{2.3}$$

where the score variable is not standardized to unit standard deviation, and the remaining variables are defined above. $\tilde{\beta}$ can therefore be interpreted as an elasticity. The coefficients of interest are β and $\tilde{\beta}$, which we expect to be positive, i.e. a high score is predicted to be associated with more forward citations. This is the hypothesis that will be tested. The regressions above will also be estimated using the private economic value of patents as dependent variable, another measure of importance of a patent. Note that the citations are counted from the grant date of a patent, yet the scores are computed using the filing year of a patent as the time of reference for determining the technological landscape relevant at the time of innovation.

2.3.3.2 Firm-level outcomes

At the firm level, the aim is to estimate the effect of filing patents with a high score on firm-level outcomes in the following years. We choose the filing date as opposed to the grant date to define the event, because a firm may start using a patented invention before the patent is granted. The firm-level outcomes of interest are output, profits, capital and employment. Patent-level information must be aggregated at the firm level. One firm may file several patents in a given year — especially in this sample of large firms. Patents with a high score are typically rare and filed only by few firm-year observations. We define *high-score patents* as patents whose scores are in the top 5% of the overall score distribution net of year fixed effects.[17] The variable of interest at the firm level is a dummy denoted $D_{fi,t}$ that takes value 1 if firm f from industry i — defined as the 3-digit SIC code — filed such a patent in year t. We consider a dummy as opposed to the number of top patents since firm-year observations that file a top patent typically represent a small fraction of the observations, and observations with more than one top patent are even rarer (as noted by Kelly et al., 2018, p. 29).

The generic specification of interest at the firm level closely follows that in Kelly et al. (2018, equation 19) and reads:

$$\log\left[\frac{1}{\mid h\mid}\sum_{\tau=1}^{h} Y_{fi,t+\tau}\right] - \log Y_{fi,t} = \beta_h D_{fi,t} + \gamma \mathbf{Z}_{fi,t} + \mathbf{FE}_{it} + \epsilon_{fi,t+h}, \tag{2.4}$$

where $Y_{fi,t}$ is either output, profits, capital or employment of firm f from industry i in year t, $\mathbf{Z}_{fi,t}$ is a vector of controls and $D_{fi,t}$ is defined above.[18] The left-hand side of the regression is the growth rate of the average of outcome variable Y between t and $t+h$ relative to its value in year t.[19] Following Kelly et al. (2018), we consider horizons

---

[17] The results are also estimated for a threshold of 1% for robustness, and are available upon request.

[18] This specification is also similar to equation (12) in Kogan et al. (2017).

[19] The coefficients for $h < 0$ are therefore the growth rate between $t+h$ and h, taking t as the base year. A negative β_h hence indicates a positive growth rate if $h < 0$.

between -5 and +10, i.e. from 5 years before and up to 10 years after the filing date.[20] In all these regressions, we control for the log of total assets, the age of the firm since its entry into the data, a dummy for whether the firm filed a patent in that year, the log of (1+) the number of filed patents in that year, a dummy for whether the firm is in the top percentile(s) in terms of number of patents filed in that year, the share of top patents among filed patents in that year (for that firm), as well as in the stock of patents up to $t-1$ (for that firm). The lag of the level of the outcome variable — whose growth rate is the dependent variable — is also included in all regressions. By including year-industry fixed effects, the aim to is compare how a firm that files a top patent fares relative to other firms in that industry, at that time. Standard errors are clustered two-way at the firm and year levels.[21]

The coefficients of interest are the sequence of β_h. This specification allows to test for the absence of pre-trends: $\beta_h \, \forall h < 0$ should not be significantly different from 0, otherwise that firm may already be on a different trend before the filing date, and this may be unrelated to the invention. This dynamic specification also allows to study the effect of filing a top patent over time, as it is not obvious when — if at all — the effects should be apparent.

It should be noted that the dependent variable in equation (2.4) is very smooth over time for a given firm, i.e. it will vary little and especially so in the later years. This is because the average profits over longer horizons is not very sensitive to the last data point added to the average. Furthermore, a stance needs to be taken relative to missing data points as it is unclear what should the value of the average growth rate be if a value (or more) is missing at some point of time within the horizon of interest, which is not uncommon given the unbalanced nature of the panel. For robustness and transparency, we drop any observation which has a missing or more at any given horizon. This changes the composition of the sample over h whereby the effects for large h are estimated based on firms that survive at least for that long. However, this way of proceeding guarantees that firms with missing data are not artificially kept in the sample when they are in fact not observed.[22] As a last robustness check, we

---

[20] Kelly et al. (2018) also use the filing year as the event date.

[21] Autocorrelation-consistent standard errors were also used (in conjunction with clustering) for robustness as the dependent variable is likely to be auto-correlated over time. The resulting standard errors are usually much smaller. We chose to err on the side of caution and report the main results with the regular clustered standard errors.

[22] The results generally look stronger and the sequence of β_h is smoother when keeping observations with missing data points, but we feel it is partly artificially driven.

also estimate (2.4) with the growth rate of firms outcomes over the whole horizon as dependent variable, namely $\log Y_{fi,t+h} - \log Y_{fi,t}$ as in Kogan et al. (2017).

2.4 Results

In this section, we go through our results. We first document shifts in economy-wide and field-specific topics of innovation, and study the success of patents and firms anticipating these shifts. Second, we show that patents with high widening scores perform particularly well and so do the firms filing them, although there is evidence that these firms were already outperforming their peers prior to filing these patents. Last, we study long-term trends in our innovation scores in an attempt in understand whether ideas have become harder to find over time.

2.4.1 Macro and field-specific shifts in innovation

We first attempt to visualise economy-wide (macro) and IPC-specific shifts in innovation and to see whether patents anticipating these shifts have higher citations and benefit the filing firms. For this purpose, we use the concept of centroids which was introduced and discussed in sections 2.3.1 and 2.3.2.

2.4.1.1 Visual intuition

First recall that a centroid is the mean vector of a set of patents. This set of patents is a subset of Z. For a macro centroid, this subset contains all patents in the economy filed in a range of years. For an IPC centroid, it contains all patents filed in a range of years within a given IPC code. Put differently, centroids are imaginary mean patents at a given time, either of the whole economy or of a specific field. We are interested in the change in centroids over time.

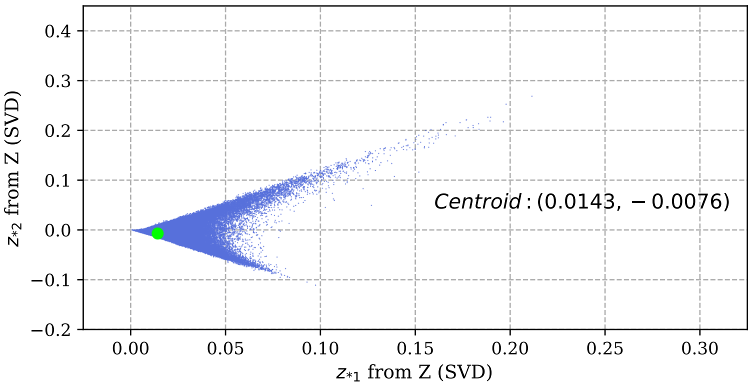
Economy-wide shifts To illustrate the exercise intuitively, consider the movement of the mean of the first two components of the tSVD between periods 1976-1985 and 2006-2015, as depicted in figure 2.2. For graphical illustration we only use patents filed at the two time extremities of our sample, but we use the whole sample in regressions. As already noted in section 2.3.1.2 the spatial information contained in the two first components is very limited.[23] Figure 2.2 shows a move in centroid vectors. Al-

---

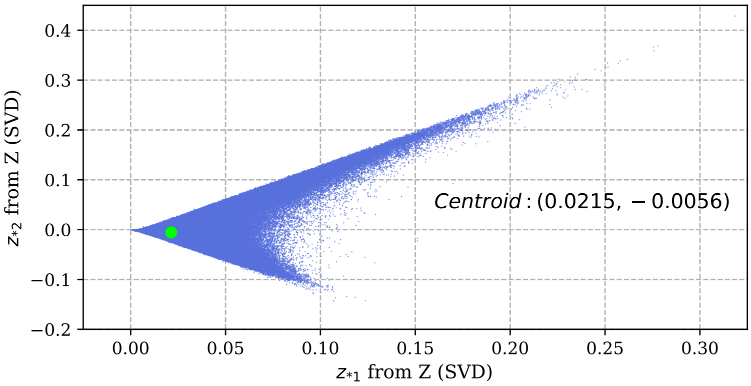
[23] We cannot use the t-SNE vectors to compute centroids as t-SNE is a tool for visualisation only and the distances between clusters are distorted.

FIGURE 2.2: ECONOMY-WIDE (MACRO) CENTROIDS (FIRST TWO SVD COMPONENTS)

(A) All patents filed from 1976 to 1985



(B) All patents filed from 2006 to 2015



Note: each dot represents a patent, whose coordinates are the two first components (columns) from Z obtained via tSVD. Panel (A) plots each patent filed over 1976-1985, and panel (B) each patent filed over 2006-2015.

though it appears tiny visually, it hopefully carries the intuition of what such changes in centroids mean.

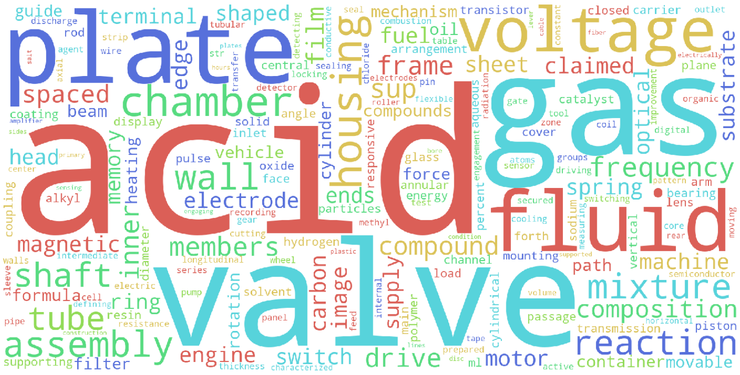
Since relying on the first two dimensions of Z does not offer much insight, we exploit a great particularity of centroids, namely that we can look “into” their language content. We compute the same centroids using the corresponding rows in the dtm (in which columns are words), which yields mean word frequencies.[24] Recomputing the same two centroids based on the dtm suggests significant centroid movements over time. Figure 2.3 illustrates that whereas the economy-wide centroid over 1976-1985 prominently features words from chemistry such as *acid* or *gas*, the centroid over 2006-2015 shows words such as *memory*, *semiconductor*, or *network*. The movement of the imaginary mean patent for the whole economy intuitively depicts the development of the computing age.

---

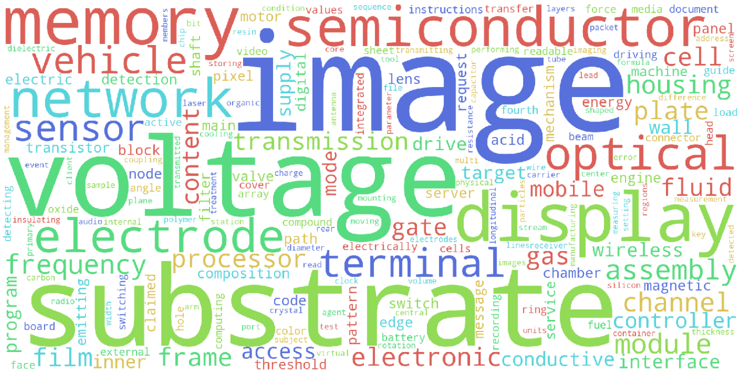
[24] Hence we can switch from numerical points in our 300-dimensional space to the words underlying in these points (subject to some approximation error).

FIGURE 2.3: WORD CONTENT OF ECONOMY-WIDE CENTROIDS

(A) Words in 1976 to 1985 macro centroid



(B) Words in 2006 to 2015 macro centroid



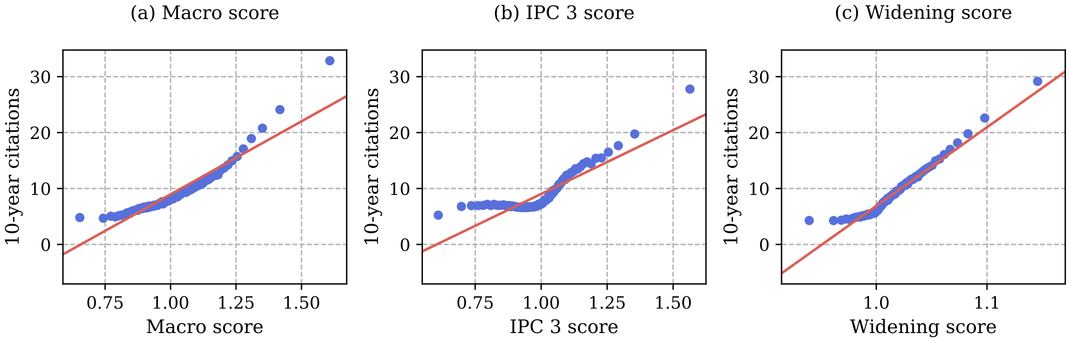
Note: the size of a word is proportional to its frequency. Words frequencies are obtained by averaging columns of the dtm over the subset of rows corresponding to the patents of interest: in panel (A), all patents filed in 1976-1985; in panel (B), all patents filed in 2006-2015.

Field-specific shifts The same methodology allows to look into shifts in innovation topics at an arbitrarily disaggregated IPC code level. First note that for many IPC 3 codes (i.e. IPC categories at the 3-digit level) the words associated with the centroids are surprisingly stable, suggesting that popular innovation topics in these fields remained stable over time.[25] We illustrate the movement in centroids for two IPC 3 fields with relatively clear changes: “H04: Electric communication technique” in figure 2.B.3, and “C01: Inorganic chemistry” in figure 2.B.4. In H04, the movement in centroid content immediately reveals the emergence of wireless communication technologies. For instance, the word *network* was hardly mentioned before and *telephone* became *mobile*. In C01 the words used indicate that chemistry patents with acids lost in importance whereas *carbon* and *gas* rose in popularity within this sub-field. Note that IPC 3 codes are still relatively aggregated. The same concept of centroid visual-

---

[25] In general, world cloud visualisations provide a good first idea. However, font size does not only relate to frequency but also e.g. to the method used to position large amount of words in a rectangular shape. Depending on the random seed chosen they can give somewhat different impressions.

FIGURE 2.4: SCORES VERSUS CITATIONS



Note: binned scatter of scores against 10-year forward citations. Number of bins in each plot is 100. Binned scatter plots with varying number of bins are provided in figures 2.B.7, 2.B.8 and 2.B.9, and scatter plots of the raw data are provided in figure 2.B.10.

isation can be used at higher levels of disaggregation. Centroids are very handy for a quick visualisation of shifts in innovation topics over time, but do not allow any kind of quantitative analysis. The next section links our centroid scores to patent- and firm-level performance measures.

2.4.1.2 Regressions

The existence of macro and field-specific shifts suggests that the timing and content of firms' innovative output may matter for how successful their patents are, and ultimately for their economic performance. Specifically, a firm that anticipates or starts a major shift in the technological landscape may reap the benefits of being among the first to do so. This idea applies both at the macro level — an invention in the IT sector in the 1990s may be very influential — and at the IPC level within a narrowly defined technological field. How far a patent is to the relevant centroid when it is filed and how close it is to the new centroid over the 10 subsequent years may be indicative of how topical a patent is. In the following sections, we study whether such patents have higher citations, and whether they are associated with better performance at the firm level. The scores in this section are the centroid-based scores as described in detail in section 2.3.2. The *macro scores* are computed with backward and forward spaces that contain all patents (their mean patent is the macro centroid) and the *IPC 3 scores* are computed with backward and forward spaces that only contain patents of the same IPC 3 as that of the patent for which the score is computed (the mean of this set of patents is an IPC centroid).

TABLE 2.4: 10-YEAR CITATIONS AND MACRO SCORE

	Dependent variable: 10-year citations					
	Whole sample			*Until 2000*		
	(1)	(2)	(3)	(4)	(5)	(6)
Macro score, standardized	3.976*** (0.000)	4.425*** (0.000)	3.873*** (0.000)	4.517*** (0.000)	4.474*** (0.000)	4.012*** (0.000)
Constant	-17.16*** (0.000)			-19.60*** (0.000)		
Year FE		✓	✓		✓	✓
IPC 3 FE		✓			✓	
Firm FE			✓			✓
Adjusted R^2	0.043	0.079	0.147	0.086	0.122	0.215
Within R^2		0.028	0.023		0.042	0.041
Observations	2,896,300	2,896,013	1,152,830	1,545,952	1,545,675	612,934

Note: *: p<0.1, **: p<0.05, ***: p<0.01. P-values from standard errors clustered at the filing year level in parenthesis.

Citations Both macro and IPC 3 scores are positively associated with patent citations, as depicted in figure 2.4.[26] Patenting an invention far from the centroid at the time of filing yet close to the new centroid 10 years later is associated with higher citations. To formalize the results, equation (2.2) is estimated using both the macro and IPC 3 scores. The results can be found in columns (1), (2) and (3) of tables 2.4 and 2.5, respectively. The relation between scores and citations is statistically significant and economically large. In the specification including filing year and IPC 3 fixed effects, a one standard deviation increase in the macro score is associated with an increase of 4 citations, more than half of the mean citations (see table 2.1). Similarly, a one standard deviation increase in the IPC 3 score is associated with an increase of 3 citations — half of the mean citations in the sample.[27] Higher IPC 3 scores are also associated with a higher economic value as measured by Stoffman et al. (2019), whereas this is not true for macro scores — see tables 2.A.4 and 2.A.5 in the appendix. All the results above remain qualitatively unchanged using citations 5 or 15 years after grant date.

It should be noted that the association between citations and scores is somewhat weaker in the last decade of our sample. Columns (4), (5) and (6) of tables 2.4 and 2.5 present the results of the same regressions as above, limiting the sample to the years up to 2000. Although the coefficients have similar magnitudes, the within R-squared are significantly higher. It is unclear why this is the case. Tentatively, it could be due

---

[26] Since binned scatter plots sometimes hide heterogeneity in the raw data, binned scatter plots with varying number of bins are provided in figures 2.B.7 and 2.B.8 for the macro and IPC 3 scores, and scatter plots of the raw data are provided in figure 2.B.10.

[27] The results using specification (2.3) can be found in the appendix, in tables 2.A.2 and 2.A.3. The results are similar.

TABLE 2.5: 10-YEAR CITATIONS AND IPC 3 SCORE

	Dependent variable: 10-year citations					
	Whole sample			*Until 2000*		
	(1)	(2)	(3)	(4)	(5)	(6)
IPC 3 score, standardized	3.071***	2.999***	2.790***	3.584***	3.417***	2.881***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-13.78***			-16.27***		
	(0.000)			(0.000)		
Year FE		✓	✓		✓	✓
IPC 3 FE		✓			✓	
Firm FE			✓			✓
Adjusted R^2	0.025	0.075	0.143	0.048	0.124	0.205
Within R^2		0.024	0.018		0.047	0.030
Observations	2,745,260	2,745,260	1,118,094	1,448,555	1,448,555	590,788

Note: *: p<0.1, **: p<0.05, ***: p<0.01. P-values from standard errors clustered at the filing year level in parenthesis.

to the increase in the number of patents filed in the later years, which results in a more crowded space in the vector representation and prevents our method from identifying meaningful centroids. This issue will be further discussed in section 2.5.

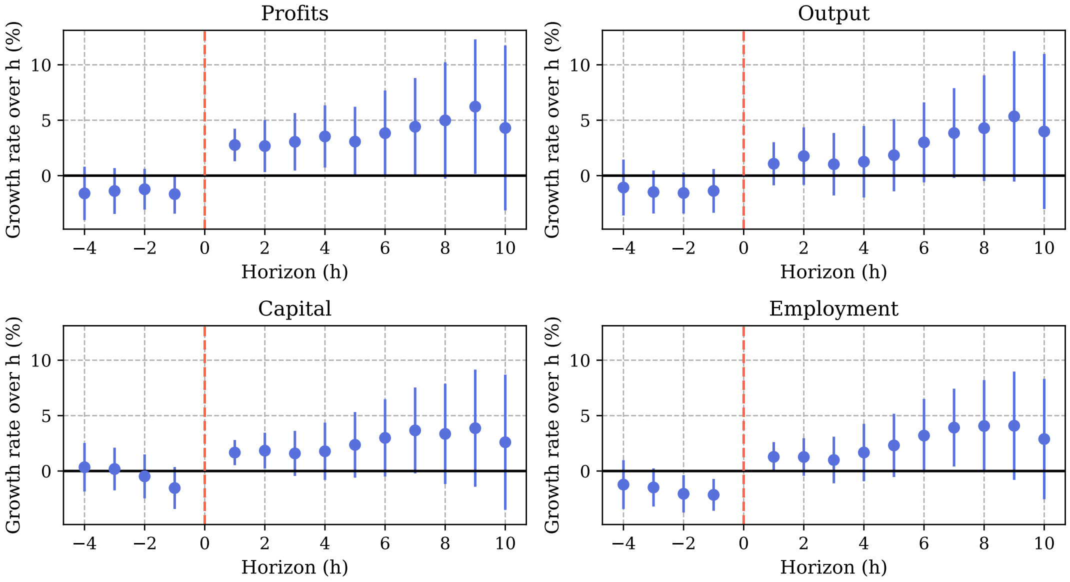
It is interesting that the relationship with citations is strong both at the macro and the IPC 3 levels. At the macro level, it is likely to reflect the advantage of being an early innovator in a field that will become important for the economy as a whole. Given the illustrations of the previous section, it seems that the macro scores are likely to capture the IT revolution, a field that was near non-existent in the 1980s and that came to dominate the technological landscape in the 1990s and 2000s. Indeed, most of the patents with high macro scores are from IT-related sectors. This is also true for IPC 3 scores, which perhaps indicates that IT is the sector whose centroid moved the most over time.

Firm-level performance A firm that files a patent with a high score may benefit from such an invention. In order to test whether this is the case, equation (2.4) is estimated for both the macro and the IPC 3 scores. The treatment variable is a dummy that indicates whether a firm filed a patent in the top 5% of the chosen score distribution, which we call a high-score patent.[28] The results are shown graphically in figures 2.5 and 2.6, which depict the sequence of estimated β_h for profits, output, capital and

---

[28] As mentioned in section 2.3.3, we consider top patents in the distribution of scores removing year fixed effects, which means that high score patents are those with the highest scores among the patents filed in the same year. We also ran the regressions using the top patents in the overall distribution without removing year fixed effects, and the effects are similar. The results are also similar when classifying the top 1% patents as high-score patents.

FIGURE 2.5: TOP MACRO PATENTS AND FIRMS DYNAMICS



Note: estimates from equation (2.4) using the macro score to qualify top patents. 95% confidence intervals are depicted. The coefficient is the growth rate between year $t + h$ and t using t as the base year, so for negative h a negative coefficient implies a positive growth rate between $t + h$ and t.

employment.[29]

The results suggest that filing a top patent in terms of IPC 3 scores is generally associated with higher subsequent growth in output and capital for the filing firm relative to other firms in that industry and year, although the effects are not strongly significant. Profits seem to be on an upward trajectory even before the event date. These pretrends suggest that those firms that file a high-score patent are *already* growing faster than other firms *before* the innovation happens. It is therefore impossible to identify the causal effect of patenting a high-score invention on profits. Instead, our method seems to identify firms which are generally innovative. Since innovative output is highly persistent, it is perhaps not surprising that high-score patents are produced by firms that have a consistently high level of innovation and that grow faster than other firms on average. Filing a top patent in terms of the macro score does not result in any strongly statistically significant effect on profits, output, capital or employment for the filing firm.

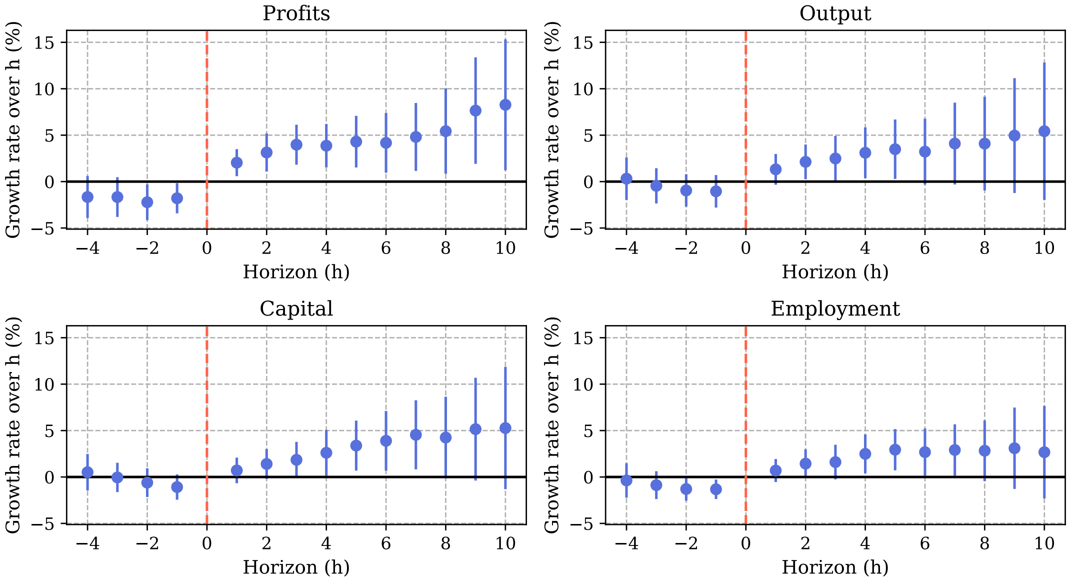
2.4.2 Widening existing ideas

We now use our geometrical approach to define a group of patents that widened knowledge by venturing into unexplored parts of the technological space which sub-

---

[29] The results estimated using the alternative specification of (2.4) with $\log Y_{fi,t+h} - \log Y_{fi,t}$ as dependent variable can be found in figures 2.B.11 and 2.B.12.

FIGURE 2.6: TOP IPC 3 PATENTS AND FIRMS DYNAMICS



Note: estimates from equation (2.4) using the IPC 3 score to qualify top patents. 95% confidence intervals are depicted. The coefficient is the growth rate between year $t + h$ and t using t as the base year, so for negative h a negative coefficient implies a positive growth rate between $t + h$ and t.

TABLE 2.6: ONE DIGIT IPC CODES (IPC SECTIONS)

Code	Name
A	Human necessities
B	Performing operations; transporting
C	Chemistry; metallurgy
D	Textiles; paper
E	Fixed constructions
F	Mechanical engineering; lighting; heating; weapons; blasting
G	Physics
H	Electricity

sequently became important.

2.4.2.1 Visual intuition

We start by providing our definition of widening patents. A *widening* patent is a patent which is dissimilar to its closest neighbours in the set of patents filed in the past 10 years, and similar to its closest neighbours in the set of patents filed in the following 10 years. In other word, a widening patent enters a part of space with few existing patents and this region is subsequently populated in the following years.

While our centroid-based scores allow us to study economy-wide and within-field shifts, they are ill-suited to identify widening patents if the neighbours of a patent belong to different IPC fields. In that case, the IPC-specific centroid may not represent the relevant technologies most related to the patent of interest. Figure 2.B.2 in section 2.3.1.2 already suggested that there may be considerable heterogeneity in terms of patents' IPC within neighbourhoods. To consolidate this argument, consider panel (A) of figure 2.7, which depicts patents filed from 1976 to 1985 based on the t-SNE algorithm introduced in section 2.3.1. Each dot represents a patent and is coloured according to the IPC section (IPC 1) to which it belongs — a list thereof can be found in table 2.6. Note that importantly, the algorithm has no information on the IPC code a patent belongs to when computing the patent vectors. A few striking patterns emerge. First, patents in chemistry (C) electronics and physics (G and H) are concentrated in relatively separated and homogeneous clouds. In other words, a patent from IPC C, G or H is mostly surrounded by patents from the same field. Patents from other IPC sections such as Human Necessities (A), Textiles (D), or Fixed Constructions (E) are much less clustered and seem to spread across many topic areas of the language space. Due to this heterogeneity in clusters, our widening score is better suited to capture widening patents.

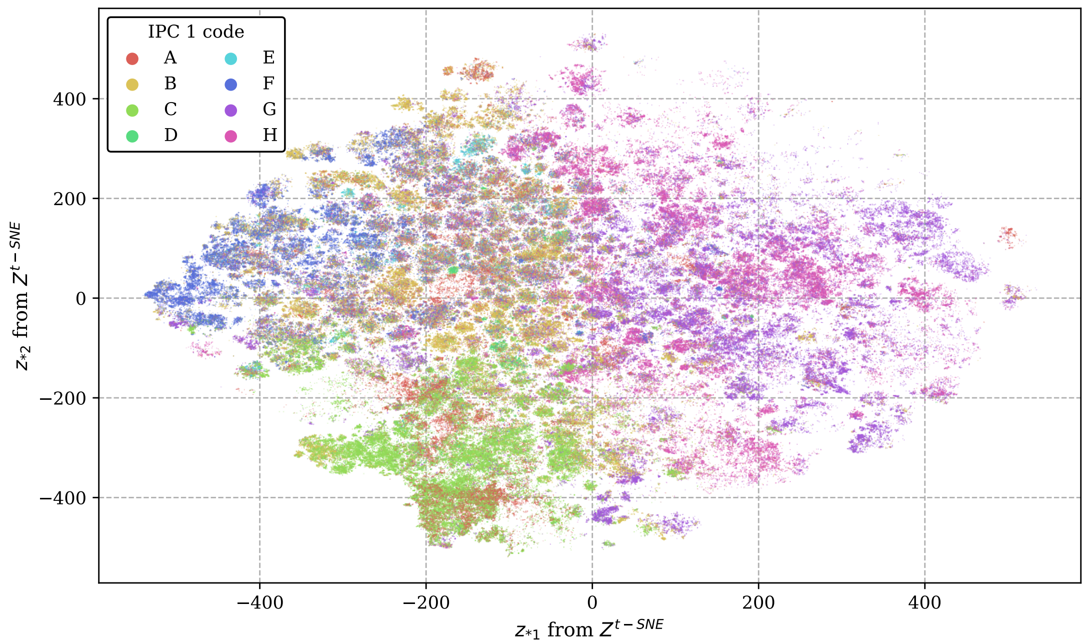
The t-SNE visualisation offers other interesting insights into the data when comparing patents filed at the two time extremities of our sample in panel (A) and (B) of figure 2.7. The areas of space for Physics (G) and Electricity (H) in 2006 to 2015 are very crowded whereas they were almost empty back in 1976-1985, in line with what we would expect from the IT revolution: whole new fields were created. We can already surmise that many widening patents will originate in these fields, as will be confirmed in section 2.5. The two snapshots of 1976-1985 and 2006-2015 also suggest that some clusters appear and disappear over time. For example, a sub-field of chemistry strongly linked with human necessities appeared around coordinates $(z_{*1} = -300, z_{*2} = -200)$ while another sub-field of chemistry at coordinates $(z_{*1} = -180, z_{*2} = -400)$ disappeared.

2.4.2.2 Regressions

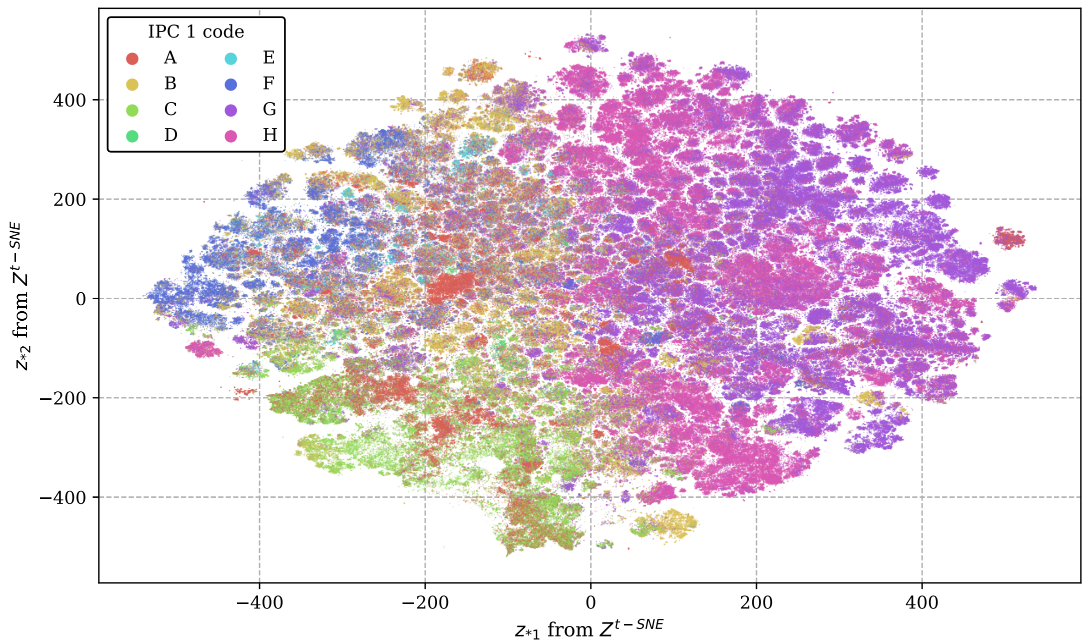
Armed with the widening scores computed as described in section 2.3.2, we analyse whether high-score patents — which we call widening patents — are successful patents in terms of citations and private value, and whether the firms from which the inventions originate benefit and perform better relative to their peers. The widening scores allow for a neater narrative and a finer analysis than the centroid scores, where the score of a patent was obtained by comparing its content to economy-wide or sector-specific *average* content at different points in time. A patent with a high

FIGURE 2.7: IPC-COLOURED T-SNE REPRESENTATIONS

(A) IPC 1-coloured t-SNE representation of patents from 1976 to 1985



(B) IPC 1-coloured t-SNE representation of patents from 2006 to 2015



Note: scatter plots of the two columns of $Z^{\text{t-SNE}}$. Each dot is a patent, coloured by the IPC 1 category to which it belongs.

widening score is a patent that was distant to its closest neighbours when filed and that got new close neighbours in the subsequent 10 years. The neighbours need not be from the same sector or field of technology.

Citations High widening scores are associated with more citations: a one standard deviation increase in the score is associated with 3 to 4 additional citations over 10

TABLE 2.7: 10-YEAR CITATIONS AND WIDENING SCORE

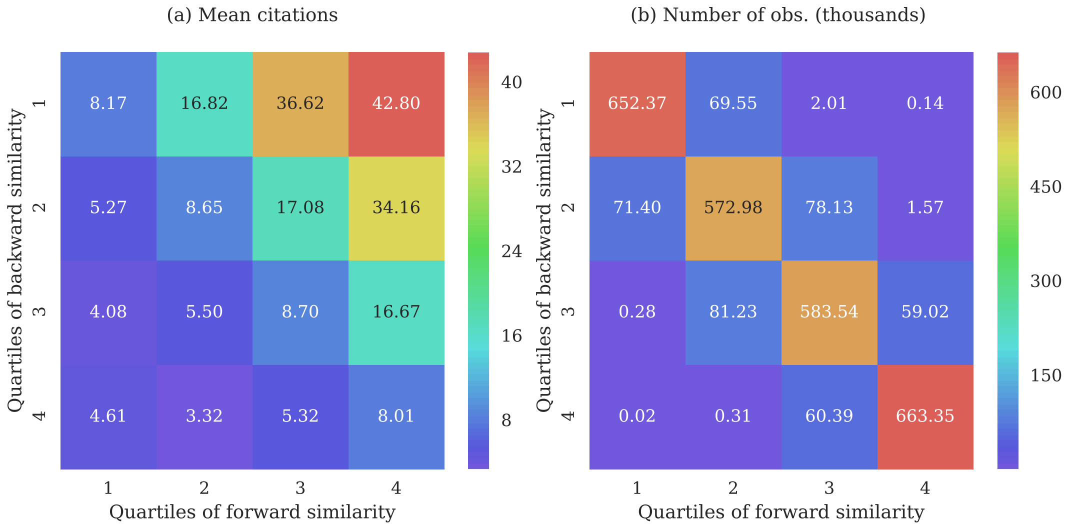
	Dependent variable: 10-year citations					
	Whole sample			*Until 2000*		
	(1)	(2)	(3)	(4)	(5)	(6)
Widening core, standardized	4.026***	3.395***	3.495***	4.334***	3.488***	3.407***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-133.9***			-145.1***		
	(0.000)			(0.000)		
Year FE		✓	✓		✓	✓
IPC 3 FE		✓			✓	
Firm FE			✓			✓
Adjusted R^2	0.044	0.075	0.148	0.067	0.119	0.212
Within R^2		0.024	0.024		0.038	0.038
Observations	2,896,300	2,896,013	1,152,830	1,545,952	1,545,675	612,934

Note: *: p<0.1, **: p<0.05, ***: p<0.01. P-values from standard errors clustered at the filing year level in parenthesis.

years — an increase representing about 50% to 60% of the mean citations over 10 years in the sample. The results can be found in table 2.7 — and table 2.A.6 for the specification in logs. Both the magnitudes and goodness of fit are similar to the results obtained using macro and IPC 3 scores, which is somewhat surprising since the definitions of the scores differ markedly. The relationship between scores is discussed in section 2.5.1. Widening patents are also associated with higher private values — as reported in table 2.A.7 — and the correlation is much higher than in the case of IPC 3 scores, but unstable across specifications using different sets of fixed effects. All the results above remain qualitatively unchanged using citations 5 or 15 years after grant date. Note again the decrease in the goodness-of-fit of these regressions in the years after 2000 — columns (1)-(3) versus (4)-(6) in tables 2.7 and 2.A.6.

A high score could in principle be the result of a high forward neighbourhood similarity and a low backward neighbourhood similarity — our definition of widening patents — but could also be driven by a very low backward neighbourhood similarity and a low forward neighbourhood similarity. In that case, the patent would be very dissimilar at the time of filing, and yet remain fairly dissimilar 10 years later — although less so. It would be difficult to argue that such a patent is a widening innovation. To confirm that the positive association between citations and widening scores is driven by wideners, figure 2.8 reports the average citations for patents in different quintiles of the distributions of backward and forward neighbourhood similarities. Citations increase as a patent is dissimilar in the backward space and similar in the forward space, which matches our narrative nicely.

FIGURE 2.8: MEAN CITATIONS PER DECILE OF BACKWARD AND FORWARD WIDENING SIMILARITY



Note: Panel (a) displays the mean citations of patents whose backward and forward neighbourhood similarities fall in the corresponding combinations of quartiles of the distributions of scores, and panel (b) reports the underlying number of observations. A high value of the quartile means higher similarity. The values on the y-axis are swapped (i.e. they range from 4 to 1).

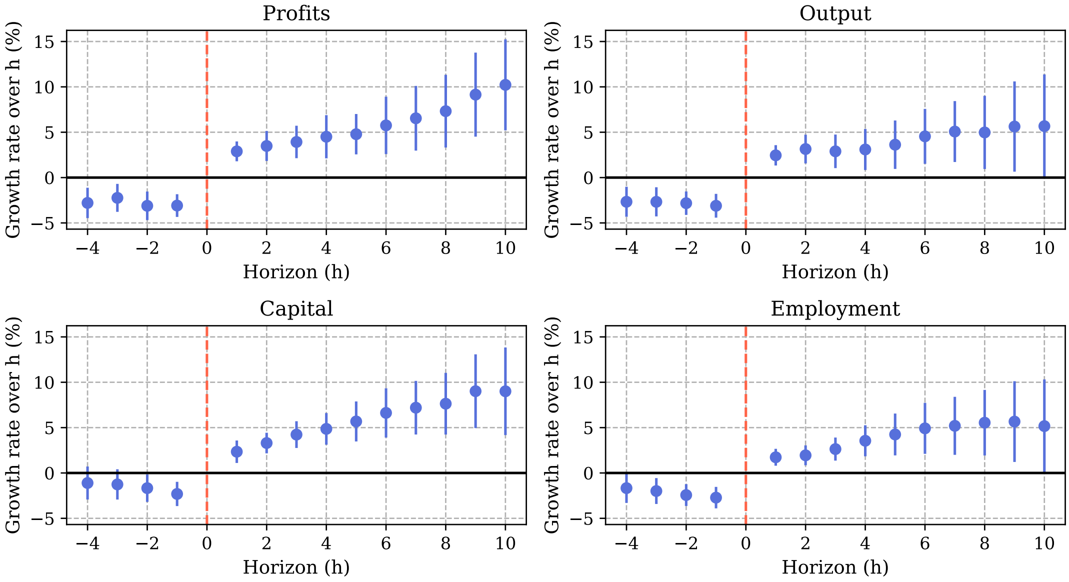
Firm-level performance The results can be found in figure 2.9.[30] Firms that file a high-score patent see their profits, output and capital grow faster than their peers within the same industry and years. There is however very strong evidence of existing pre-trends, as suggested by significantly negative coefficients before the event date (the filing of the high-score patent). Those firms were on a differential trend before the invention came out. As mentioned earlier, this renders causal statements about the effect of filing a high-score patent impossible. It seems that we identify fast-growing firms that innovate and continue to grow fast relative to other firms after the patent has been filed. It is however reassuring that the effects seem to be stronger than when high-score patents are identified using macro and IPC 3 scores since the widening scores are identifying *widening* patents — which may be rarer inventions — whereas the other two scores merely identify patents that anticipated a general shift in innovation. However, the firms with high widening scores are often IT firms. Table 2.A.1 lists the firms with most top patents in the distribution of widening scores (i.e. patents in top 5 percentiles of the score distribution net of year fixed effects). They are all IT-related.

It is interesting to note that differential pre-trends can also be observed when considering the filing of top patents in terms of citations and private value. Specifically,

---

[30] The results estimated using the alternative specification of (2.4) with $\log Y_{fi,t+h} - \log Y_{fi,t}$ as dependent variable can be found in figure 2.B.13

FIGURE 2.9: WIDENING PATENTS AND FIRMS DYNAMICS



Note: estimates from equation (2.4) using the widening score to qualify top patents. 95% confidence intervals are depicted. The coefficient is the growth rate between year $t+h$ and t using t as the base year, so for negative h a negative coefficient implies a positive growth rate between $t+h$ and t.

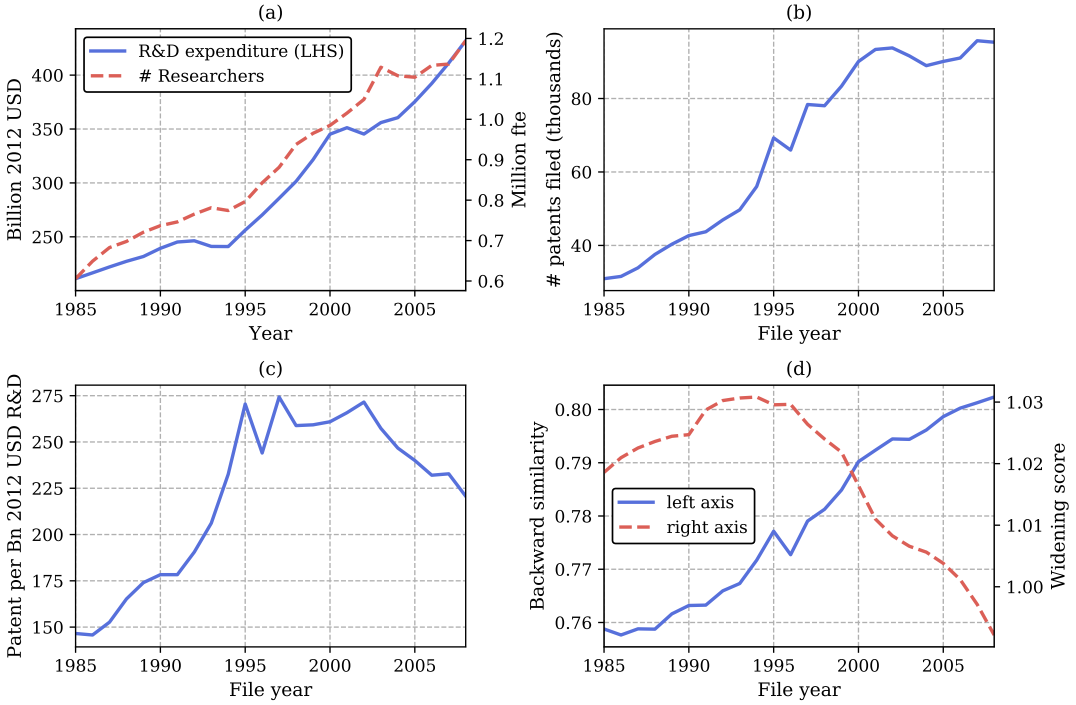
estimating equation (2.4) classifying top patents as being in the top 5% of the private value distribution or in the top 0.1% of the citations distributions yield pre-trends in both cases (especially marked in the case of private value) as reported in figures 2.B.17 and 2.B.18. This problem in estimating the causal effect of top patents — irrespective of the criteria used to qualify them — seems to be a recurrent issue.

2.4.3 Have ideas become harder to find?

In this section, we apply our methodology to a question hotly debated in the growth literature, namely whether ideas are becoming harder to find over time. There has been recent evidence that while research effort has increased substantially over the last decades, research productivity has declined in tandem (Gordon, 2017; Bloom, Jones, Van Reenen and Webb, 2020). Because the relationship between ideas and patents may vary over time, counting patent per dollar of R&D expenditure may be misleading (Lanjouw and Schankerman, 2004). We attempt to understand whether ideas have gotten harder to get over our sample period based on our widening scores.

The premise of this exercise is that we can proxy the novelty of ideas with our widening score and its components, specifically the backward similarity that measures how different a patent is to the stock of existing patents. We study the evolution of these measures over time and across technological fields to investigate whether ideas

FIGURE 2.10: AGGREGATE TRENDS IN THE US



Note: panel (a): Gross domestic spending on R&D in billion 2012 USD and number of researchers in full time equivalents (source: OECD series *GDEXPRD* and *RESEARCHER*); panel (b): number of patents filed by US entities; panel (c): number of patents filed per billion USD spent on R&D (contemporaneous); panel (d): average backward neighbourhood similarity to 100 closest neighbours and widening score across patents filed by US entities.

are getting harder to get, taking into account research effort as measured by R&D expenditures. The fundamental idea is that if patents become less innovative over time conditional on research effort, ideas may have become harder to get.

2.4.3.1 Aggregate trends

As a motivation, figure 2.10 plots economy-wide trends in the US over the period 1985-2008. Several striking patterns emerge. First, research effort as measured by R&D spending and research workers have consistently increased over the years. The number of patents filed yearly by US entities also steeply increased up until the 2000s when it plateaued. Number of patents per dollar of R&D increased over 1985-1995, stabilized afterwards, and started to decline after 2000 (this pattern is similar for patents per researcher).[31] If every patent was an idea of constant quality, one would tentatively infer that ideas might have become harder to find sometime between 1995 and 2000, resulting in less innovative output per research input.

---

[31] This calculation is only motivational, as R&D expenditures probably result in patents in the following years and not in the current year. The graphs plots number of patents per R&D billion dollars contemporaneously spent.

However, ideas may not be of constant quality and a patent may protect an invention which is more or less novel and impactful. Panel (d) in figure 2.10 plots the average backward neighbourhood similarity of patents to their 100 closest neighbours — the denominator in our widening score — as well as the average widening score over time. The average backward neighbourhood similarity consistently increases over time, suggesting that the content of patents becomes more similar over time. The average widening score increases up until 1995 and sharply decreases afterwards. Combined, these trends suggest that novel ideas may have gotten harder to find over time after 1995. The next sections dissect the data across innovation fields to offer further support to this conclusion. A discussion of potential caveats of our approach follows in section 2.5.

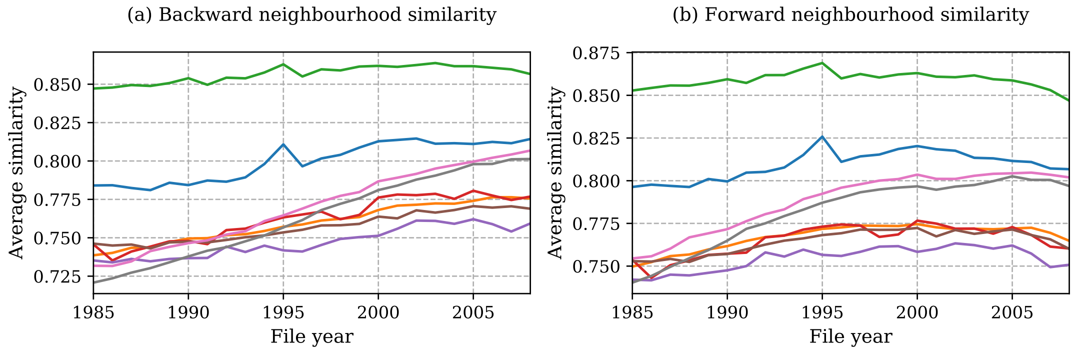
2.4.3.2 Comparison across fields

These aggregate trends may hide insightful heterogeneity as some fields saw the birth of numerous innovative and original ideas (e.g. the IT-related fields) whereas other fields did not. We would like to explore whether it may have been easier to find ideas in fields with high widening scores relative to other fields. The average scores over time per IPC 1 codes are plotted in figure 2.11 for all patents and in figure 2.B.22 for patents linked to Compustat firms, in which case data on R&D is also available at the IPC level. There is heterogeneity in levels and trends across fields. First, IPC G and H (physics and electricity, where the IT revolution mostly originated) display higher widening scores on average, and steeper increases over time in both backward and forward neighbourhood similarities, as well as in the number of filed patents. The decline in the widening scores post-1995 is also the most pronounced. Second, IPC A (human necessities) and C (Chemistry and metallurgy) both display high and stable backward and forward neighbourhood similarities.[32] Third, the remaining fields appear to exhibit similar patterns: low and mildly increasing backward for forward neighbourhood similarities. Forth and perhaps most importantly, all sectors experience a similar hump-shaped patterns in their average widening scores over time, peaking in the early 1990s and decreasing afterwards, driven by a flattening in the forward neighbourhood similarity relative to the backward neighbourhood similarity.

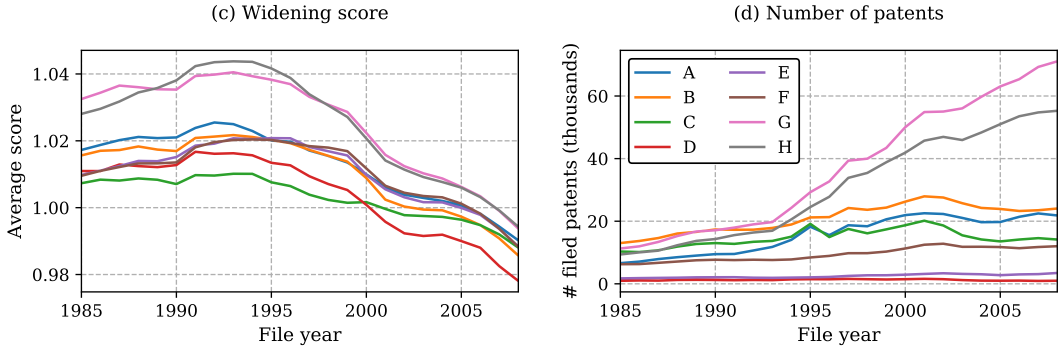
---

[32] The spike in both the average backward and forward neighbourhood similarities in 1995 are likely due to a large inflow of patents of similar contents which appear in both the backward and forward spaces. There is indeed a corresponding spike in the number of patents filed in 1995.

FIGURE 2.11: TRENDS BY IPC 1


(a) Backward neighbourhood similarity

(b) Forward neighbourhood similarity


(c) Widening score

(d) Number of patents

Note: panel (a): average backward neighbourhood similarity by IPC 1 over time; panel (b): average forward neighbourhood similarity by IPC 1 over time; panel (c): average widening score by IPC 1 over time; panel (d): number of patents filed by IPC 1 over time. Sample: all patents.

This last pattern may suggest that impactful ideas are getting harder to find: inventions become more similar to those in the past, and spawn less related innovations in the future. It is however surprising that this pattern is present *in each* IPC 1 fields. This could be consistent with new IT-related words being newly used across all IPC sections — and not only in patents of IPC sections G and H.[33] From figure 2.B.22, one can see that R&D expenditures at the firm level kept rising. Patents per dollar of R&D are relatively stable for most fields, except A and C where it decreased significantly.[34]

To ease the reading of the data, we divide the sample into two groups of technological fields according to their average widening score over the sample period. We denote fields with high and low average widening score *higher-* and *lower-innovation fields*, respectively. We restrict the sample to IPC 3 codes with at least 10,000 patents (this represents 95% of the observations) and obtain two groups of 31 IPC 3 fields each. The trends for each fields group are displayed in figure 2.12 (and figure 2.B.23

---

[33] A similar pattern can be seen in figure 3 of the updated version of Kelly et al. (2018) — the working paper dated February 2020. However, they do not dissect the data by IPC 1, but by the percentiles of the widening score.

[34] We refrain from overly interpreting the decrease in the number of patents filed by Compustat firms over years 2003-2008 as this may be a decrease in the number of firm-patent matches as opposed to an actual decrease in filing. This decrease in the matched patents is also visible in panel (b) of figure 2.B.1.

for patents linked to Compustat firms, with R&D figures). Higher innovation fields — dominated by IPC 3 from sections B, G and H — are characterized by an increasing backward neighbourhood similarity and a forward neighbourhood similarity that first increases up to 1995 and then flattens, resulting is a strongly hump-shaped widening score. Patents filed exploded over the sample period. Lower-innovation fields — predominantly from sections C and F — exhibit similar yet much less pronounced patterns, and flows of patents filed increased only slightly over the sample period. These trends are consistent with the interpretation that innovation has been relatively easier in higher-innovation fields in the late 1980s and early 1995s before hitting decreasing returns as ideas became harder to get, whereas they have been relatively hard to find in lower-innovation fields throughout the sample. Taking into account R&D expenditures in the firm-matched sample, much more money is spent in higher-innovation fields, yet the number of patents per dollar of R&D expenditures plateaus around 1995. Overall, this tentatively suggests that innovation was once easier in the higher-innovation fields before becoming harder.[35]

2.5 Discussion

In this section, we discuss how the different scores relate to each other and emphasize the importance of IT-related innovation for our results. We finish with explanations of potential problems our method might be subject to.

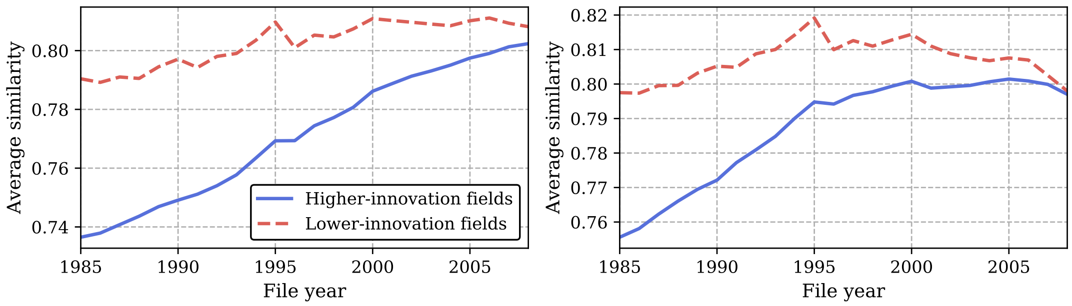
2.5.1 Relationships between scores

Recall that macro, IPC and widening scores are calculated for every patent in our sample between 1985 and 2008. As explained in section 2.3.2, the scores are the ratio of a forward similarity and a backward similarity, which are essentially similarities to a reference point capturing the existing technological states at different point in time. A patent dissimilar to existing patents at the time of filing and similar 10 years later will have a high score. Table 2.8 reports some descriptive statistics. First note that both the

---

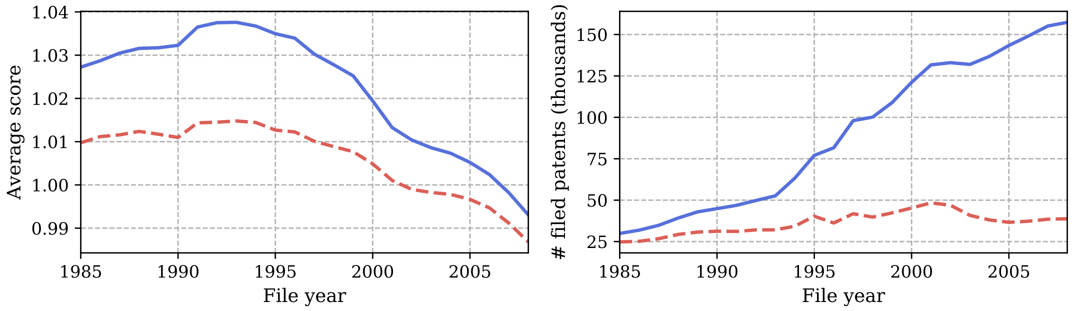
[35] As a last tentative piece of evidence in support of the claim that innovation may have been easier in higher-innovation fields, we show anecdotal evidence that returns to improvements of existing technologies may be higher in lower-innovation fields. The intuition is that a successful improvement would be more rewarded in fields where ideas are hard to find and innovation is low. We compare patents containing at least 5 keywords associated with the notion of improvement to the rest of the patents and find that the difference in average citations between the former and the latter is larger in lower-innovation fields relative to higher-innovation fields, as shown in table 2.A.11. However, these words are more likely to appear in longer documents, and length itself is positively correlated with citations. This could therefore be purely mechanical.

FIGURE 2.12: TRENDS BY IPC 3 GROUP



(a) Backward neighbourhood similarity

(b) Forward neighbourhood similarity



(c) Widening score

(d) Number of patents

Note: panel (a): average backward neighbourhood similarity by IPC 3 group over time; panel (b): average forward neighbourhood similarity by IPC 3 group over time; panel (c): average widening score by IPC 3 group over time; panel (d): number of patents filed by IPC 3 group over time. IPC 3 fields are divided into two groups according to their average widening scores. Sample: all patents.

TABLE 2.8: DESCRIPTIVE STATISTICS OF INNOVATION SCORES

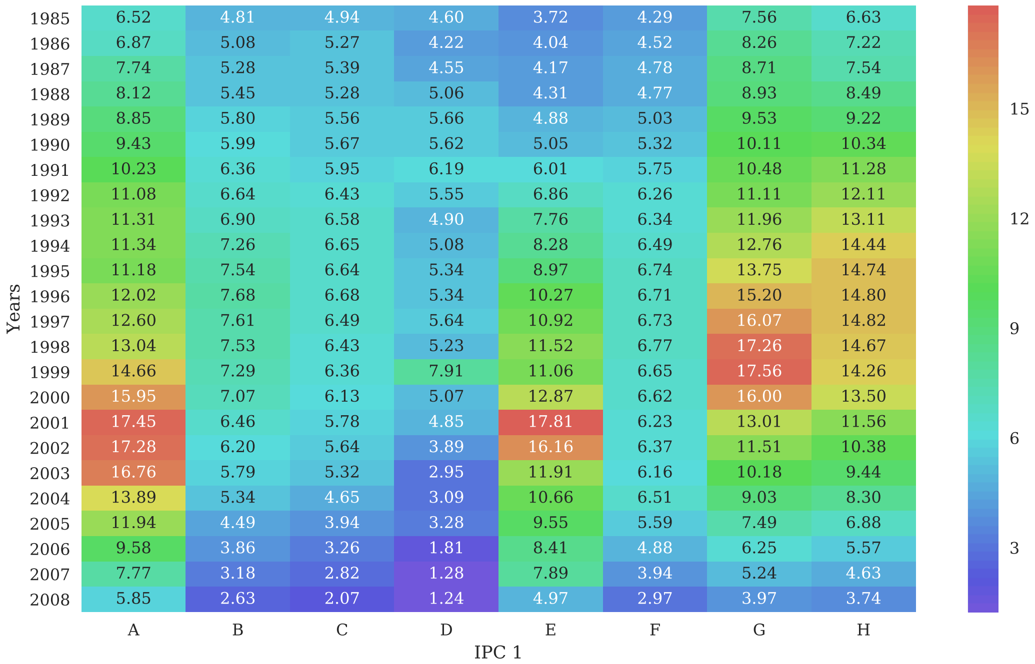
	N	Mean	Sd	Min	p25	Median	p75	p90	Max
Macro score	2,896,300	.994	.152	.093	.887	.966	1.08	1.19	3.06
IPC 3 score	2,745,260	.997	.135	.0899	.927	.993	1.06	1.14	4.74
Widening score	2,896,300	1.01	.0286	.551	.998	1.01	1.02	1.05	2.36

mean and the median scores are close to 1, which means that on average the position of a patent relative to its neighbours in the backward and forward spaces is very similar. Forward and backward similarities are indeed very highly correlated — see figure 2.B.5. The scores are positively correlated with each other with correlations ranging from 0.33 to 0.46. A correlogram of the scores containing the histograms of each score distribution and scatter plots of the scores against each other can be found in figure 2.B.6. Even though all these scores capture similar notions, they differ markedly from one another. In light of these low correlations, in is somewhat surprising that all scores exhibit similar correlations with citations (cf. tables 2.4, 2.5 and 2.7).

2.5.2 The importance of IT

An approach which identifies innovative patents as those that were dissimilar to past language but similar to future language will tend to give highest scores to

FIGURE 2.13: CITATIONS PER IPC



Years	A	B	C	D	E	F	G	H
1985	6.52	4.81	4.94	4.60	3.72	4.29	7.56	6.63
1986	6.87	5.08	5.27	4.22	4.04	4.52	8.26	7.22
1987	7.74	5.28	5.39	4.55	4.17	4.78	8.71	7.54
1988	8.12	5.45	5.28	5.06	4.31	4.77	8.93	8.49
1989	8.85	5.80	5.56	5.66	4.88	5.03	9.53	9.22
1990	9.43	5.99	5.67	5.62	5.05	5.32	10.11	10.34
1991	10.23	6.36	5.95	6.19	6.01	5.75	10.48	11.28
1992	11.08	6.64	6.43	5.55	6.86	6.26	11.11	12.11
1993	11.31	6.90	6.58	4.90	7.76	6.34	11.96	13.11
1994	11.34	7.26	6.65	5.08	8.28	6.49	12.76	14.44
1995	11.18	7.54	6.64	5.34	8.97	6.74	13.75	14.74
1996	12.02	7.68	6.68	5.34	10.27	6.71	15.20	14.80
1997	12.60	7.61	6.49	5.64	10.92	6.73	16.07	14.82
1998	13.04	7.53	6.43	5.23	11.52	6.77	17.26	14.67
1999	14.66	7.29	6.36	7.91	11.06	6.65	17.56	14.26
2000	15.95	7.07	6.13	5.07	12.87	6.62	16.00	13.50
2001	17.45	6.46	5.78	4.85	17.81	6.23	13.01	11.56
2002	17.28	6.20	5.64	3.89	16.16	6.37	11.51	10.38
2003	16.76	5.79	5.32	2.95	11.91	6.16	10.18	9.44
2004	13.89	5.34	4.65	3.09	10.66	6.51	9.03	8.30
2005	11.94	4.49	3.94	3.28	9.55	5.59	7.49	6.88
2006	9.58	3.86	3.26	1.81	8.41	4.88	6.25	5.57
2007	7.77	3.18	2.82	1.28	7.89	3.94	5.24	4.63
2008	5.85	2.63	2.07	1.24	4.97	2.97	3.97	3.74

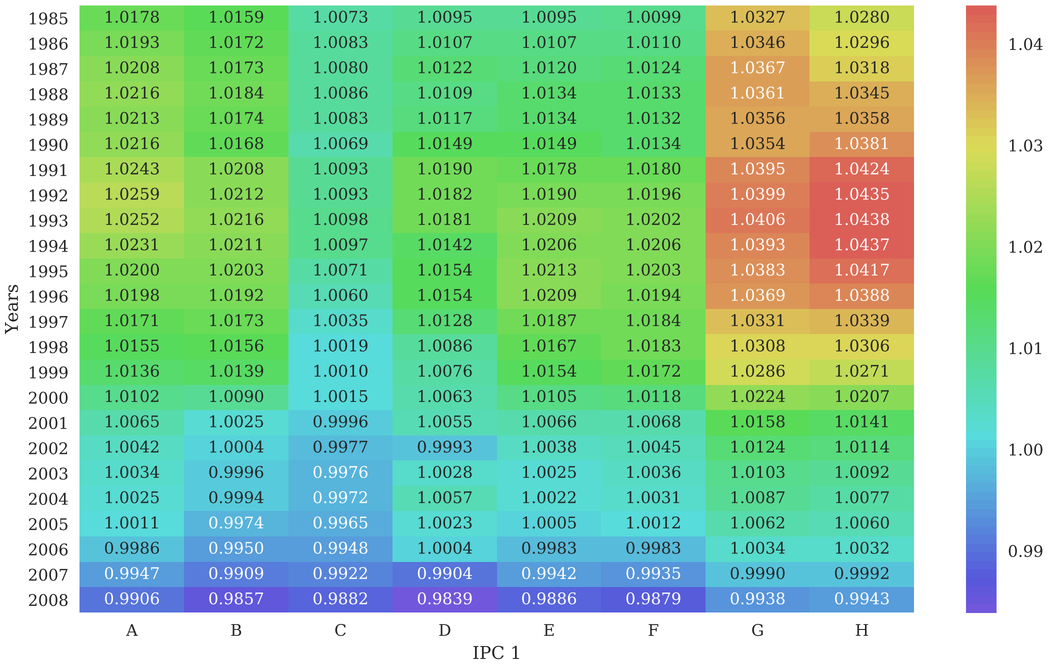
Note: each cell is the average number of citations within 10 years of the filing year, by year and IPC section (IPC 1). Cells are colored such that higher values have hotter colors.

those patents which were in empty parts of the space when they got released. Our t-SNE-driven intuition suggests that patents with high scores will come in large proportions from electronics and physics fields as these areas have been the emptiest in the early years of our sample. The heatmaps in figures 2.13 and 2.14 confirm this intuition. The former shows average 10-year citations per year-IPC 1 combination and the latter displays the corresponding widening scores. As expected, our widening score is highest in IT-related areas, which is where the space must have been the emptiest.

Figures 2.B.19, 2.B.20, and 2.B.21 show the same for macro and IPC 3 scores, as well as depict the number of patents per cell. In the case of macro centroids, IT patents in the 1990s have even higher scores relative to other patents, reflecting the move of the entire economy towards IT technologies over the time span we consider. In the case of IPC 3 scores, average score values are more balanced as each IPC code has its own reference point. It seems that methods such as the ones discussed here — as well as in other related papers, in all likelihood — identify the IT revolution.

One possible worry is that our results are solely driven by patents relating to the IT sector. At the macro level, IT is one important sector that was nascent in the 1980s, and became central in the subsequent decades. At the IPC level, IPC G and H —

FIGURE 2.14: WIDENING SCORES



Years	A	B	C	D	E	F	G	H
1985	1.0178	1.0159	1.0073	1.0095	1.0095	1.0099	1.0327	1.0280
1986	1.0193	1.0172	1.0083	1.0107	1.0107	1.0110	1.0346	1.0296
1987	1.0208	1.0173	1.0080	1.0122	1.0120	1.0124	1.0367	1.0318
1988	1.0216	1.0184	1.0086	1.0109	1.0134	1.0133	1.0361	1.0345
1989	1.0213	1.0174	1.0083	1.0117	1.0134	1.0132	1.0356	1.0358
1990	1.0216	1.0168	1.0069	1.0149	1.0149	1.0134	1.0354	1.0381
1991	1.0243	1.0208	1.0093	1.0190	1.0178	1.0180	1.0395	1.0424
1992	1.0259	1.0212	1.0093	1.0182	1.0190	1.0196	1.0399	1.0435
1993	1.0252	1.0216	1.0098	1.0181	1.0209	1.0202	1.0406	1.0438
1994	1.0231	1.0211	1.0097	1.0142	1.0206	1.0206	1.0393	1.0437
1995	1.0200	1.0203	1.0071	1.0154	1.0213	1.0203	1.0383	1.0417
1996	1.0198	1.0192	1.0060	1.0154	1.0209	1.0194	1.0369	1.0388
1997	1.0171	1.0173	1.0035	1.0128	1.0187	1.0184	1.0331	1.0339
1998	1.0155	1.0156	1.0019	1.0086	1.0167	1.0183	1.0308	1.0306
1999	1.0136	1.0139	1.0010	1.0076	1.0154	1.0172	1.0286	1.0271
2000	1.0102	1.0090	1.0015	1.0063	1.0105	1.0118	1.0224	1.0207
2001	1.0065	1.0025	0.9996	1.0055	1.0066	1.0068	1.0158	1.0141
2002	1.0042	1.0004	0.9977	0.9993	1.0038	1.0045	1.0124	1.0114
2003	1.0034	0.9996	0.9976	1.0028	1.0025	1.0036	1.0103	1.0092
2004	1.0025	0.9994	0.9972	1.0057	1.0022	1.0031	1.0087	1.0077
2005	1.0011	0.9974	0.9965	1.0023	1.0005	1.0012	1.0062	1.0060
2006	0.9986	0.9950	0.9948	1.0004	0.9983	0.9983	1.0034	1.0032
2007	0.9947	0.9909	0.9922	0.9904	0.9942	0.9935	0.9990	0.9992
2008	0.9906	0.9857	0.9882	0.9839	0.9886	0.9879	0.9938	0.9943

Note: each cell is the average widening score, by year and IPC section (IPC 1). Cells are colored such that higher values have hotter colors.

Physics and Electricity, the IPC seemingly most related to IT — account for over 50% of granted patents in our sample. Regarding *widening* patents, we also find that large IT firms disproportionately file high-score patents — see table 2.A.1. It is therefore natural to wonder whether the results are robust to excluding patents from these technological areas. We do so by dropping all patents from IPC codes G and H once the scores have been computed, and re-estimate all our results.[36] In general, the results are qualitatively unchanged, although in most cases the magnitudes of the effects are smaller — they sometimes as much as halve. The patent-level regression results can be found in tables 2.A.8, 2.A.9 and 2.A.10 for macro, IPC 3 and widening scores, respectively. The decrease in magnitude is strongest for the macro scores: in the citations regressions, both coefficients and r-squared decrease by around 50%. For results based on IPC 3 and widening scores, the decrease in magnitude is smaller. At the firm level, the estimated effects of filing a top patent are generally smaller and less significant, but qualitatively similar as can be seen in figures 2.B.14, 2.B.15 and 2.B.16.[37]

One marked difference is the relationship between private value and scores, which

---

[36] We do not, however, re-estimate the vector representation of patents dropping IPC codes G and H, i.e. we use the same scores as in the main text, simply omitting patents from those IPC codes. Omitting IT patents before computing patent representations would result in severely biased patent vectors.

[37] The differential pre-trend issue is also less severe, but this is probably only due to the decrease in significance across all coefficients and not a sign that differential pre-trends are less worrying in this subsample.

is most cases becomes *negative*. It is unclear why that is so: it seems that the positive relationship between scores and private value was solely driven by patents from IPC codes G and H — it is not the case that these patents have systematically higher values than those in other fields, however. Overall, it seems that IT alone is indeed responsible for some, yet not all of the results presented.

2.5.3 Construction of the scores

The widening scores are built using the proximity of patents to their 100 most similar neighbours in space. This cutoff is arbitrary and entails a trade-off. On the one hand, taking the distance to more patents, e.g. 500, might dilute the information that we attempt to capture on the proximity of the closest inventions, by taking the average distance over too many patents. On the other hand, choosing too few patents, e.g. 10, may also give a misleading idea of the position of a patent relative to its neighbours — e.g. if the 10 are very close and filed by the same firm, and the 11th is really far. We chose 100 as it seemed to be a good compromise, but the results are similar when using alternative thresholds.[38] Furthermore, it is not obvious that this threshold must remain constant over time. Since the space becomes increasingly crowded over time as the volume of patents increases, the closest 100 to a patent in 1985 may be further away than the closest 100 to a patent in 2008 even if patent texts were completely random. An alternative could be to let the threshold change over time proportionally to the number of patents in the forward and backward spaces.

2.5.4 Potential measurement issues

There are a number of limitations to our approach relating to the construction of the scores that we would like to submit to the reader. These issues will guide our future work.

First, distances between patents in the vector space may have different meanings across fields. As an hypothetical example, two patents relating to chemistry may have many words in common yet describe completely different inventions, whereas two patents in the field of computing sharing many words may be similar inventions. If this difference were constant over time, the widening scores would not be impacted as they are the ratio of two similarities. It could be an issue if it changes over time.

---

[38] We tried 50, 100, 200 and 500.

Suppose that differences between IT-related patents are expressed in fewer and fewer words, i.e. that the texts of IT-patents become more similar over time, yet the difference in technological content remains identical. The widening score would increase over time even if the rate of innovation remains the same. A similar trend would be observed if the language used in patents converged over time and became more homogeneous irrespective of the technological content of inventions.[39]

Second, the numbers of patents present in the backward and forward spaces generally increase over time as the number of patents filed per year tends to increase.[40] Since the vector space has the same dimensions throughout — it has been built based on the document term matrix of the entire set of patents — the "density" of the space increases over time.[41] Using the 2-dimensional example, if the plane is a square of fixed dimensions, as the number of dots (patents) increases, the space between patents decreases on average — even if coordinates of the dots were random. In reality, patent vectors are not random and higher-innovation fields experience much more drastic increases in patent numbers. Even if these patents were worthless, they would use some words found in other patents within these fields, resulting in increasingly crowded backward and forward spaces. It is unclear whether this is problematic and the observed trends mechanically result from this increase in density, or whether more crowded spaces indicate that innovation is harder. Further work is required to better understand these dynamics.

Third, forward and backward neighbourhood similarities may be influenced by strategic patenting practices by large firms, who strategically file large amounts of patents of related contents in an effort to create "patent thickets" (Shapiro, 2000) characterized by numerous patents with overlapping content and owned by several firms (Noel and Schankerman, 2013). New entrants wanting to innovate face potentially complex legal battles and might be discouraged to do so, *de facto* yielding a competitive advantage to these large firms. This phenomenon seems to be particularly acute in IT-related fields. Overall, patents of very similar contents might be the result of

---

[39] This is unlikely since these words would appear in more that 15% of documents and be deleted from the abstracts when we build our document representations.

[40] The numbers of patents present in the backward and forward spaces can decrease if the flow of newly filed patents is smaller than the flow of patents exiting the spaces.

[41] The dimension-reduced matrix Z is computed with all patents from all sectors and years together, i.e. on the entire matrix X with roughly 4.6 million patents. We then extract different subsets of patent vectors from this matrix and analyse them throughout this paper. If we think of the convex hull spanned around all row vectors in Z as having a fixed size, analysing larger subsets of vectors it could mechanically decrease the distance between vectors as more of this convex hull is filled with vectors. This might then mechanically decrease innovation scores towards the end of the sample when most patents are filed.

these strategies and this will impact our measures of neighbourhood similarities.

2.6 Conclusion

In this paper, we apply methods from natural language processing to analyse the innovative content of patents based on their text. Numerical representations of patents as high-dimensional vectors allow us to compute similarities between documents at a large scale. We structure our thoughts around the idea that patents dissimilar to past inventions and similar to future ones may have anticipated or started shifts in innovation topics, making the particularly successful. Our measures of success are patent citations, private value and performance indicators of the filing firm. We find that patents which anticipate economy-wide or field-specific shifts in topical innovation are more successful, as are those that *widen* knowledge in the technological space. Causal interpretation of the effect of filing such patents on firms' outcomes is rendered difficult by the existence of differential pre-trends: firms that file such patents already outperform their competitors prior to filing the patent. We show that our results are mainly driven by the IT revolution. Last, we study trends in our measures to provide tentative evidence that ideas may have gotten harder to get over time, especially in the fields that experienced high levels of innovation over the last decades.

2.7 Bibliography

Arts, Sam, Bruno Cassiman, and Juan Carlos Gomez, "Text matching to measure patent similarity," *Strategic Management Journal*, 2018, *39* (1), 62–84.

Balsmeier, Benjamin, Mohamad Assaf, Tyler Chesebro, Gabe Fierro, Kevin Johnson, Scott Johnson, Guan-Cheng Li, Sonja Lück, Doug O'Reagan, Bill Yeh et al., "Machine learning and natural language processing on the patent corpus: Data, tools, and new measures," *Journal of Economics & Management Strategy*, 2018, *27* (3), 535–553.

Bloom, Nicholas, Charles I Jones, John Van Reenen, and Michael Webb, "Are ideas getting harder to find?," *American Economic Review*, 2020, *110* (4), 1104–44.

Bowen, Donald E, Laurent Frésard, and Gerard Hoberg, "Technological disruptiveness and the evolution of IPOs and sell-outs," 2018.

Deerwester, Scott, Susan T Dumais, George W Furnas, Thomas K Landauer, and Richard Harshman, "Indexing by latent semantic analysis," *Journal of the American society for information science*, 1990, *41* (6), 391–407.

Elman, Gerry J, "Automated Patent Examination Support—A Proposal," *Biotechnology Law Report*, 2007, *26* (5), 435–436.

Gordon, Robert J, *The rise and fall of American growth: The US standard of living since the civil war*, Vol. 70, Princeton University Press, 2017.

Hall, Bronwyn H, Adam Jaffe, and Manuel Trajtenberg, "Market value and patent citations," *RAND Journal of economics*, 2005, pp. 16–38.

Katila, Riitta and Gautam Ahuja, "Something old, something new: A longitudinal study of search behavior and new product introduction," *Academy of management journal*, 2002, *45* (6), 1183–1194.

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy, "Measuring technological innovation over the long run," Technical Report, National Bureau of Economic Research 2018.

Klette, Tor Jakob and Samuel Kortum, "Innovating firms and aggregate innovation," *Journal of political economy*, 2004, *112* (5), 986–1018.

Kline, Patrick, Neviana Petkova, Heidi Williams, and Owen Zidar, "Who profits from patents? rent-sharing at innovative firms," *The Quarterly Journal of Economics*, 2019, *134* (3), 1343–1404.

Kobak, Dmitry and Philipp Berens, "The art of using t-SNE for single-cell transcriptomics," *bioRxiv*, 2019.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman, "Technological innovation, resource allocation, and growth," *The Quarterly Journal of Economics*, 2017, *132* (2), 665–712.

Lanjouw, Jean O and Mark Schankerman, "Patent quality and research productivity: Measuring innovation with multiple indicators," *The Economic Journal*, 2004, *114* (495), 441–465.

Linderman, George C., Manas Rachh, Jeremy G. Hoskins, Stefan Steinerberger, and Yuval Kluger, "Fast interpolation-based t-SNE for improved visualization of single-cell RNA-seq data," *Nature Methods*, 2019, *16* (3), 243–245.

Lodh, Suman and Maria Rosa Battaggion, "Technological breadth and depth of knowledge in innovation: the role of mergers and acquisitions in biotech," *Industrial and Corporate Change*, 2014, *24* (2), 383–415.

Moorthy, Subba and Douglas E Polley, "Technological knowledge breadth and depth: performance impacts," *Journal of Knowledge Management*, 2010, *14* (3), 359–377.

Nicholas, Tom, "Does innovation cause stock market runups? Evidence from the great crash," *American Economic Review*, 2008, *98* (4), 1370–96.

Noel, Michael and Mark Schankerman, "Strategic patenting and software innovation," *The Journal of Industrial Economics*, 2013, *61* (3), 481–520.

Ozman, Müge, "Breadth and depth of main technology fields: an empirical investigation using patent data," *Science and Technology Policies Research Centre, Working Paper Series*, 2007, *7* (01).

Packalen, Mikko and Jay Bhattacharya, "New ideas in invention," Technical Report, National Bureau of Economic Research 2015.

Shapiro, Carl, "Navigating the patent thicket: Cross licenses, patent pools, and standard setting," *Innovation policy and the economy*, 2000, *1*, 119–150.

Stoffman, Noah, Michael Woeppel, and M Deniz Yavuz, "Small Innovators: No Risk, No Return," *Kelley School of Business Research Paper*, 2019, (19-5).

Strang, Gilbert, *Introduction to Linear Algebra*, 5 ed. 2016.

van der Maaten, Laurens and Geoffrey Hinton, "Visualizing data using t-SNE," *Journal of machine learning research*, 2008, *9* (Nov), 2579–2605.

Appendix 2.A Additional tables

TABLE 2.A.1: FIRMS WITH MOST TOP PATENTS, WIDENING SCORE

Company name	# of patents	# of top patents	top in % of all top patents	top in % of all own patents
INTL BUSINESS MACHINES CORP	71,134	7,809	5.4	11
MOTOROLA SOLUTIONS INC	17,967	3,060	2.1	17
HITACHI LTD	44,870	2,926	2	6.5
CANON INC	37,651	2,518	1.7	6.7
NEC CORP	23,083	2,381	1.6	10
HP INC	25,653	2,229	1.5	8.7
PANASONIC CORP	35,297	2,200	1.5	6.2
MICROSOFT CORP	21,169	2,187	1.5	10
AT&T CORP	8,563	2,078	1.4	24
SONY CORP	29,547	2,069	1.4	7
INTEL CORP	22,738	1,841	1.3	8.1
GENERAL ELECTRIC CO	27,283	1,498	1	5.5
LUCENT TECHNOLOGIES INC	8,646	1,205	0.83	14
TEXAS INSTRUMENTS INC	15,350	1,144	0.79	7.5
SUN MICROSYSTEMS INC	7,563	1,065	0.74	14
TELEFONAKTIEBOLAGET LM ERICS	7,078	1,056	0.73	15
XEROX HOLDINGS CORP	15,080	933	0.64	6.2
EASTMAN KODAK CO	17,524	924	0.64	5.3
GENERAL MOTORS CO	13,528	828	0.57	6.1
MICRON TECHNOLOGY INC	19,602	821	0.57	4.2

Note: list of firms owning the highest number of top patents over years 1985-2008, ranked using the widening score. The second column indicates the number of patent belonging to that firm, the third column indicates that of top patents belonging to that firm (patents whose score ranks in the top 5% of the overall score distribution, controlling for year fixed effects); the forth and fifth columns contain the fractions of top patents accruing to that firm (i) out of all the top patents and (ii) out of all the patents of that firm.

TABLE 2.A.2: 10-YEAR CITATIONS AND MACRO SCORE: LOG SPECIFICATION

	Dependent variable: Log(1+10-year citations)					
	Whole sample			*Until 2000*		
	(1)	(2)	(3)	(4)	(5)	(6)
Log(macro score)	1.991***	2.032***	1.732***	2.194***	1.976***	1.768***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	1.623***			1.828***		
	(0.000)			(0.000)		
Year FE		✓	✓		✓	✓
IPC 3 FE		✓			✓	
Firm FE			✓			✓
Adjusted R^2	0.069	0.167	0.212	0.117	0.163	0.217
Within R^2		0.040	0.038		0.048	0.052
Observations	2,896,300	2,896,013	1,152,830	1,545,952	1,545,675	612,934

Note: *: p<0.1, **: p<0.05, ***: p<0.01. P-values from standard errors clustered at the filing year level in parenthesis.

Table 2.A.3: 10-year citations and IPC 3 score: log specification

	Dependent variable: Log(1+10-year citations)					
	Whole sample			*Until 2000*		
	(1)	(2)	(3)	(4)	(5)	(6)
Log(IPC 3 score)	1.378***	1.434***	1.188***	1.572***	1.589***	1.188***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	1.618***			1.837***		
	(0.000)			(0.000)		
Year FE		✓	✓		✓	✓
IPC 3 FE		✓			✓	
Firm FE			✓			✓
Adjusted R^2	0.027	0.162	0.202	0.046	0.163	0.199
Within R^2		0.033	0.024		0.052	0.031
Observations	2,745,260	2,745,260	1,118,094	1,448,555	1,448,555	590,788

Note: *: p<0.1, **: p<0.05, ***: p<0.01. P-values from standard errors clustered at the filing year level in parenthesis.

Table 2.A.4: Private value and macro score

	Dependent variable: Log(private value)					
	Whole sample			*Until 2000*		
	(1)	(2)	(3)	(4)	(5)	(6)
Log(macro score)	-0.629**	0.298	0.209**	-0.909**	0.438	-0.0250
	(0.016)	(0.204)	(0.041)	(0.013)	(0.179)	(0.841)
Constant	1.803***			1.687***		
	(0.000)			(0.000)		
Year FE		✓	✓		✓	✓
IPC 3 FE		✓			✓	
Firm FE			✓			✓
Adjusted R^2	0.002	0.123	0.887	0.005	0.120	0.906
Within R^2		0.000	0.001		0.000	0.000
Observations	1,153,764	1,153,567	1,152,830	613,783	613,608	612,934

Note: *: p<0.1, **: p<0.05, ***: p<0.01. P-values from standard errors clustered at the filing year level in parenthesis.

Table 2.A.5: Private value and IPC 3 score

	Dependent variable: Log(private value)					
	Whole sample			*Until 2000*		
	(1)	(2)	(3)	(4)	(5)	(6)
Log(IPC 3 score)	1.067***	0.672***	0.215***	1.236***	0.975***	0.0911**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.010)
Constant	1.792***			1.662***		
	(0.000)			(0.000)		
Year FE		✓	✓		✓	✓
IPC 3 FE		✓			✓	
Firm FE			✓			✓
Adjusted R^2	0.004	0.124	0.888	0.006	0.124	0.906
Within R^2		0.002	0.001		0.004	0.000
Observations	1,119,019	1,119,019	1,118,094	591,628	591,628	590,788

Note: *: p<0.1, **: p<0.05, ***: p<0.01. P-values from standard errors clustered at the filing year level in parenthesis.

Table 2.A.6: 10-year citations and widening score: log specification

	Dependent variable: Log(1+10-year citations)					
	Whole sample			*Until 2000*		
	(1)	(2)	(3)	(4)	(5)	(6)
Log(widening score)	12.51***	8.176***	7.760***	11.27***	8.324***	7.660***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	1.425***			1.536***		
	(0.000)			(0.000)		
Year FE		✓	✓		✓	✓
IPC 3 FE		✓			✓	
Firm FE			✓			✓
Adjusted R^2	0.097	0.162	0.212	0.092	0.162	0.215
Within R^2		0.035	0.037		0.046	0.049
Observations	2,896,300	2,896,013	1,152,830	1,545,952	1,545,675	612,934

Note: *: p<0.1, **: p<0.05, ***: p<0.01. P-values from standard errors clustered at the filing year level in parenthesis.

Table 2.A.7: Private value and widening score

	Dependent variable: Log(private value)					
	Whole sample			*Until 2000*		
	(1)	(2)	(3)	(4)	(5)	(6)
Log(widening score)	-1.215	3.466***	1.252***	0.122	4.270***	0.545**
	(0.492)	(0.000)	(0.000)	(0.919)	(0.000)	(0.025)
Constant	1.812***			1.663***		
	(0.000)			(0.000)		
Year FE		✓	✓		✓	✓
IPC 3 FE		✓			✓	
Firm FE			✓			✓
Adjusted R^2	0.000	0.124	0.887	0.000	0.122	0.906
Within R^2		0.002	0.002		0.003	0.000
Observations	1,153,764	1,153,567	1,152,830	613,783	613,608	612,934

Note: *: p<0.1, **: p<0.05, ***: p<0.01. P-values from standard errors clustered at the filing year level in parenthesis.

Table 2.A.8: Citations, private value and macro score: without IT

	(1) 10-year citations	(2) Log(1+10-year citations)	(3) Private value	(4) Log(private value)
Macro score, std.	2.277***		-6.457***	
	(0.000)		(0.000)	
Log(macro score)		1.799***		-1.307***
		(0.000)		(0.000)
Year FE	✓	✓	✓	✓
IPC 3 FE	✓	✓	✓	✓
Adjusted R^2	0.067	0.120	0.086	0.118
Within R^2	0.015	0.027	0.005	0.004
Observations	1,402,076	1,402,076	431,894	431,894

Note: *: p<0.1, **: p<0.05, ***: p<0.01. P-values from standard errors clustered at the filing year level in parenthesis.

TABLE 2.A.9: CITATIONS, PRIVATE VALUE AND IPC 3 SCORE: WITHOUT IT

	(1) 10-year citations	(2) Log(1+10-year citations)	(3) Private value	(4) Log(private value)
IPC 3 score, std.	2.499***		-3.452***	
	(0.000)		(0.001)	
Log(IPC 3 score)		1.595***		-0.764***
		(0.000)		(0.000)
Year FE	✓	✓	✓	✓
IPC 3 FE	✓	✓	✓	✓
Adjusted R^2	0.072	0.125	0.081	0.116
Within R^2	0.022	0.031	0.002	0.002
Observations	1,260,360	1,260,360	400,019	400,019

Note: *: p<0.1, **: p<0.05, ***: p<0.01. P-values from standard errors clustered at the filing year level in parenthesis.

TABLE 2.A.10: CITATIONS, PRIVATE VALUE AND WIDENING SCORE: WITHOUT IT

	(1) 10-year citations	(2) Log(1+10-year citations)	(3) Private value	(4) Log(private value)
Widening score, std.	2.553***		-0.996	
	(0.000)		(0.224)	
Log(widening score)		8.388***		-0.827
		(0.000)		(0.101)
Year FE	✓	✓	✓	✓
IPC 3 FE	✓	✓	✓	✓
Adjusted R^2	0.071	0.120	0.081	0.114
Within R^2	0.020	0.027	0.000	0.000
Observations	1,402,076	1,402,076	431,894	431,894

Note: *: p<0.1, **: p<0.05, ***: p<0.01. P-values from standard errors clustered at the filing year level in parenthesis.

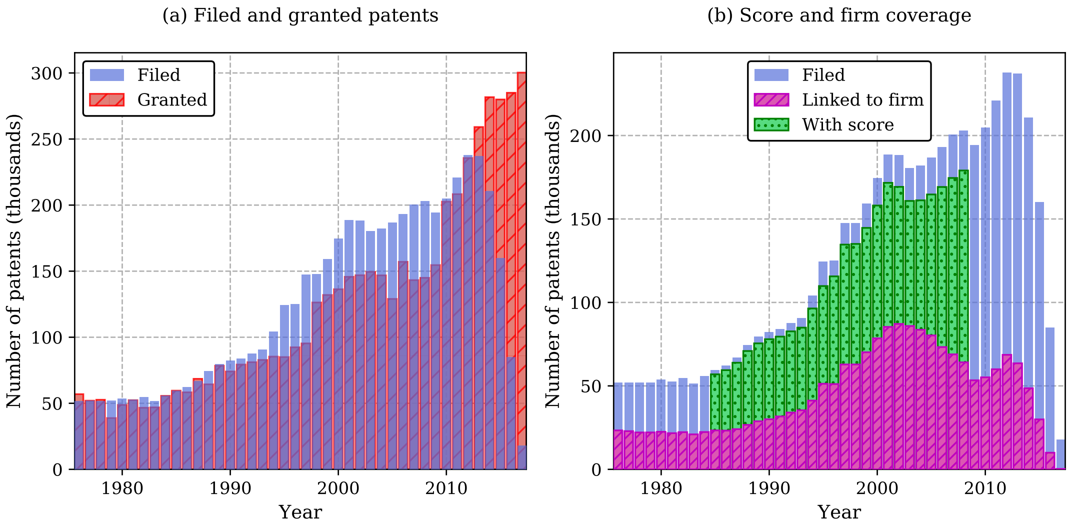
TABLE 2.A.11: RETURN TO PATENTS CONTAINING IMPROVEMENT-RELATED WORDS

	N	Length	Mean cites	Median cites	N	Length	Mean cites	Median cites
Whole sample period: 1985-2008								
	Hihger-innovation fields				Lower-innovation fields			
No key words	2,123,967	897.95	10.39	4	848,199	1,069.22	5.60	3
Key words	37,514	1,795.18	12.56	5	20,375	2,336.94	7.61	4
% change			20.89	25			35.82	33.33
Earlier years: 1985-1995								
	Hihger-innovation fields				Lower-innovation fields			
No key words	505,670	879.40	10.35	6	330,515	1,085.08	5.64	3
Key words	7,535	2,019.98	12.99	7	7,320	2,130.46	7.84	5
% change			25.41	16.67			39.03	66.67
Later years: 1996-2008								
	Hihger-innovation fields				Lower-innovation fields			
No key words	1,618,297	904.07	10.40	4	517,684	1,058.72	5.58	2
Key words	29,979	1,738.68	12.45	4	13,055	2,452.71	7.48	3
% change			19.74	0			34.07	50

Note: mean and meadian 10-year citations by innovation field groups (higher- versus lower-innovation IPC 3 fields ranked by their average widening score over the sample priod) for patents containing at least 5 key words signalling improvement versus other patents. The key words are: advance, advances, enhance, enhances, improve, improves, refine, refines, substitute and substitutes. Column *N* is the number of patents per group and *Length* is the average number of words in an abstract.

Appendix 2.B Additional figures

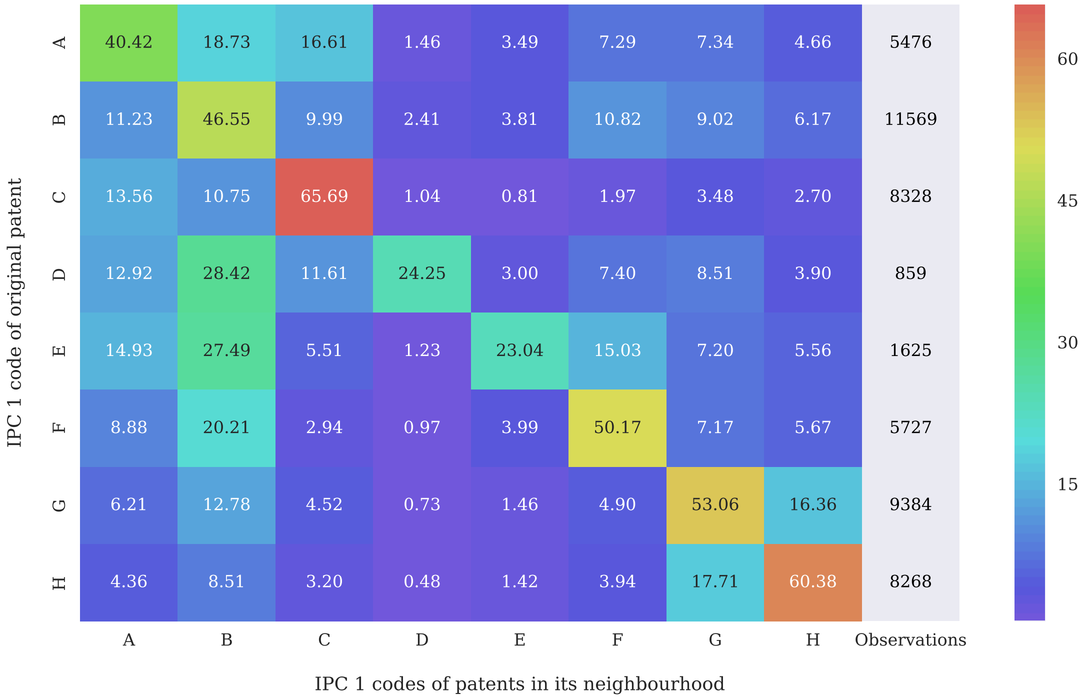
FIGURE 2.B.1: NUMBER OF PATENTS OVER TIME



Note: panel (a): number of patents filed and granted per year. Panel (b) number of patents for which a score is available, and that can be linked to firms in Compustat using the matches by Stoffman et al. (2019).

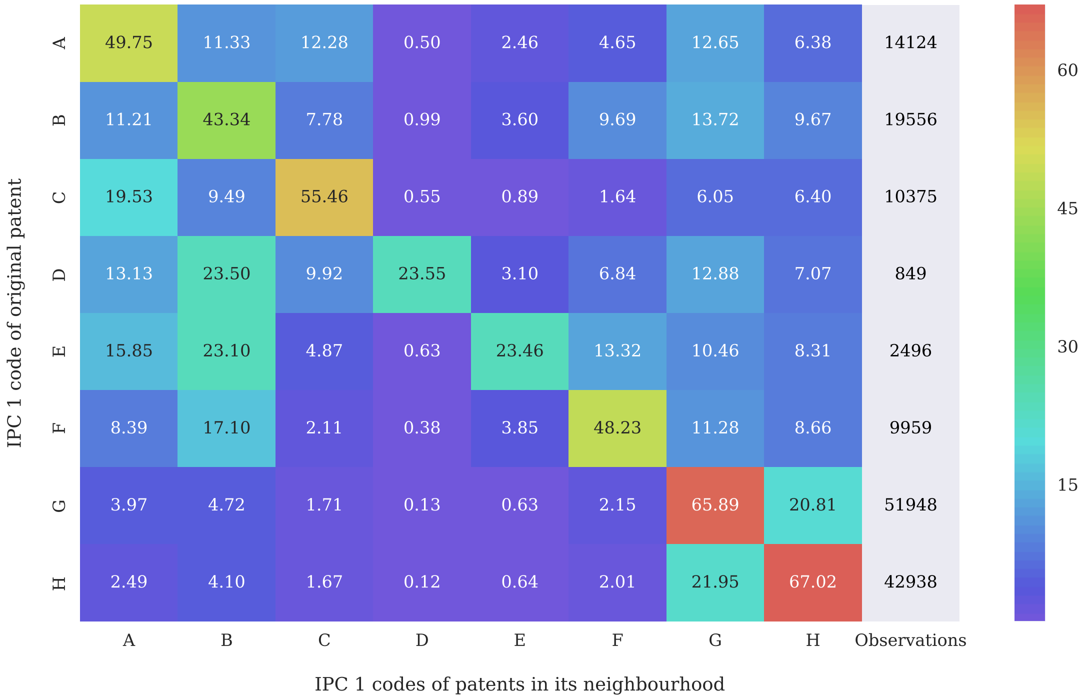
FIGURE 2.B.2: NEIGHBOURHOOD COMPOSITION BY IPC 1

(A) IPC distributions in patent neighbourhoods 1976 - 1985



	A	B	C	D	E	F	G	H	Observations
A	40.42	18.73	16.61	1.46	3.49	7.29	7.34	4.66	5476
B	11.23	46.55	9.99	2.41	3.81	10.82	9.02	6.17	11569
C	13.56	10.75	65.69	1.04	0.81	1.97	3.48	2.70	8328
D	12.92	28.42	11.61	24.25	3.00	7.40	8.51	3.90	859
E	14.93	27.49	5.51	1.23	23.04	15.03	7.20	5.56	1625
F	8.88	20.21	2.94	0.97	3.99	50.17	7.17	5.67	5727
G	6.21	12.78	4.52	0.73	1.46	4.90	53.06	16.36	9384
H	4.36	8.51	3.20	0.48	1.42	3.94	17.71	60.38	8268

(B) IPC distributions in patent neighbourhoods 2006 - 2015

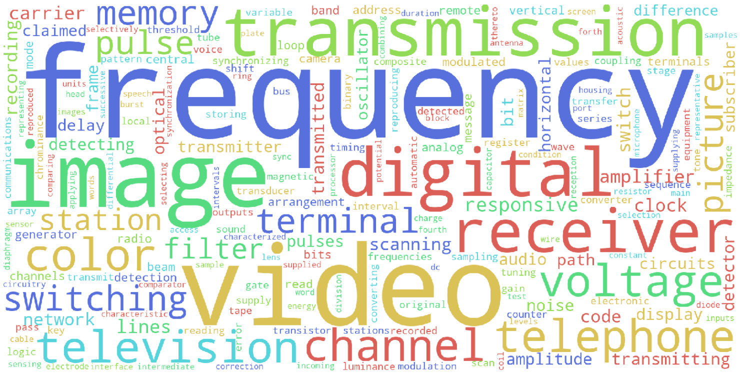


	A	B	C	D	E	F	G	H	Observations
A	49.75	11.33	12.28	0.50	2.46	4.65	12.65	6.38	14124
B	11.21	43.34	7.78	0.99	3.60	9.69	13.72	9.67	19556
C	19.53	9.49	55.46	0.55	0.89	1.64	6.05	6.40	10375
D	13.13	23.50	9.92	23.55	3.10	6.84	12.88	7.07	849
E	15.85	23.10	4.87	0.63	23.46	13.32	10.46	8.31	2496
F	8.39	17.10	2.11	0.38	3.85	48.23	11.28	8.66	9959
G	3.97	4.72	1.71	0.13	0.63	2.15	65.89	20.81	51948
H	2.49	4.10	1.67	0.12	0.64	2.01	21.95	67.02	42938

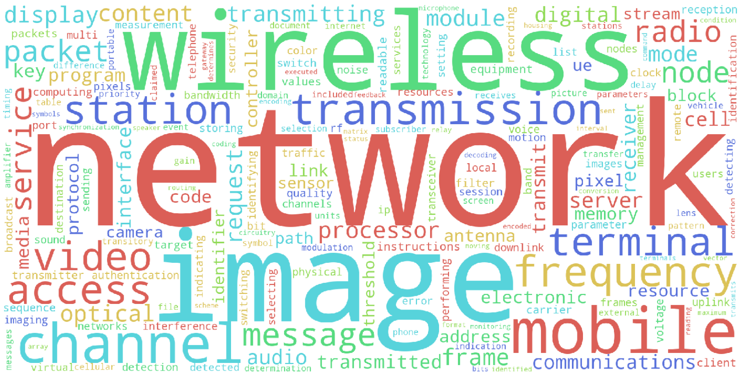
Note: for a patent belonging to the IPC 1 code in reach row, a cell contains the average share of its 100 closest neighbouring patents by IPC 1 code. Each row sums up to 100. Panel (A) is based on patents filed in the years 1976-1985 and panel (B) on patents filed in 2006-2015.

Figure 2.B.3: Word content of field-specific centroid of "H04: Electric communication technique"

(A) Words in 1976 to 1985 *H*04 IPC centroid



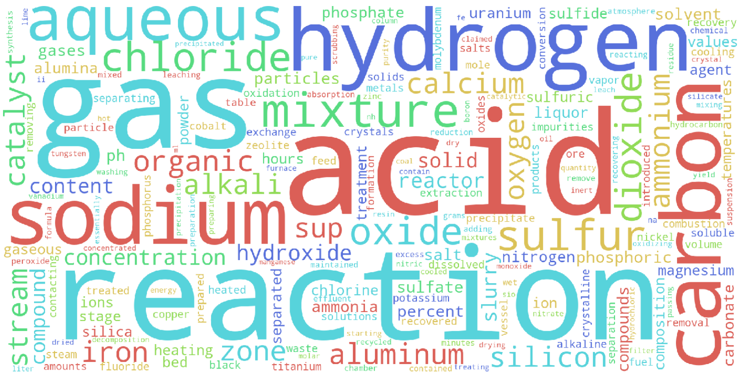
(B) Words in 2006 to 2015 *H*04 IPC centroid



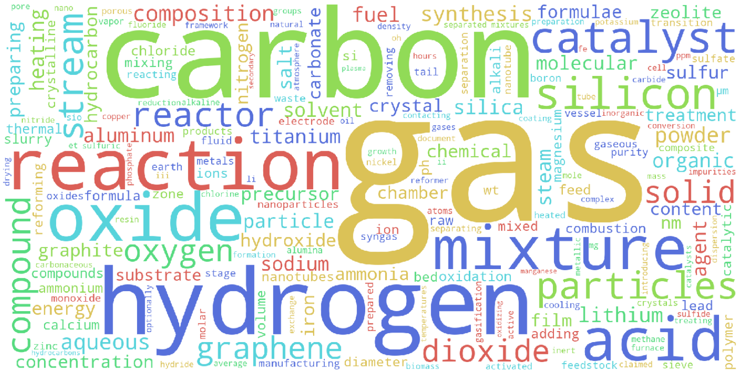
Note: the size of a word is proportional to its frequency. Words frequencies are obtained by averaging columns of the dtm over the subset of rows corresponding to the patents of interest: in panel (A), patents filed in 1976-1985 in IPC 3 H04; in panel (B), patents filed in 2006-2015 in IPC 3 H04.

FIGURE 2.B.4: WORD CONTENT OF FIELD-SPECIFIC CENTROID OF “C01: INORGANIC CHEMISTRY”

(A) Words in 1976 to 1985 C01 IPC centroid

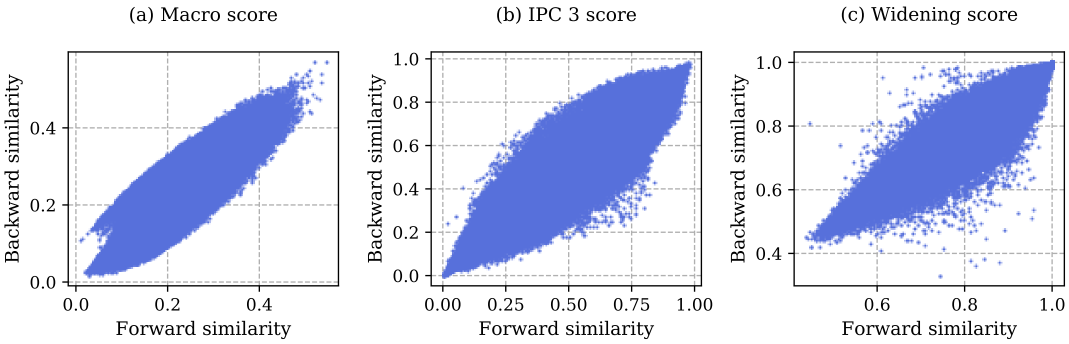


(B) Words in 2006 to 2015 C01 IPC centroid



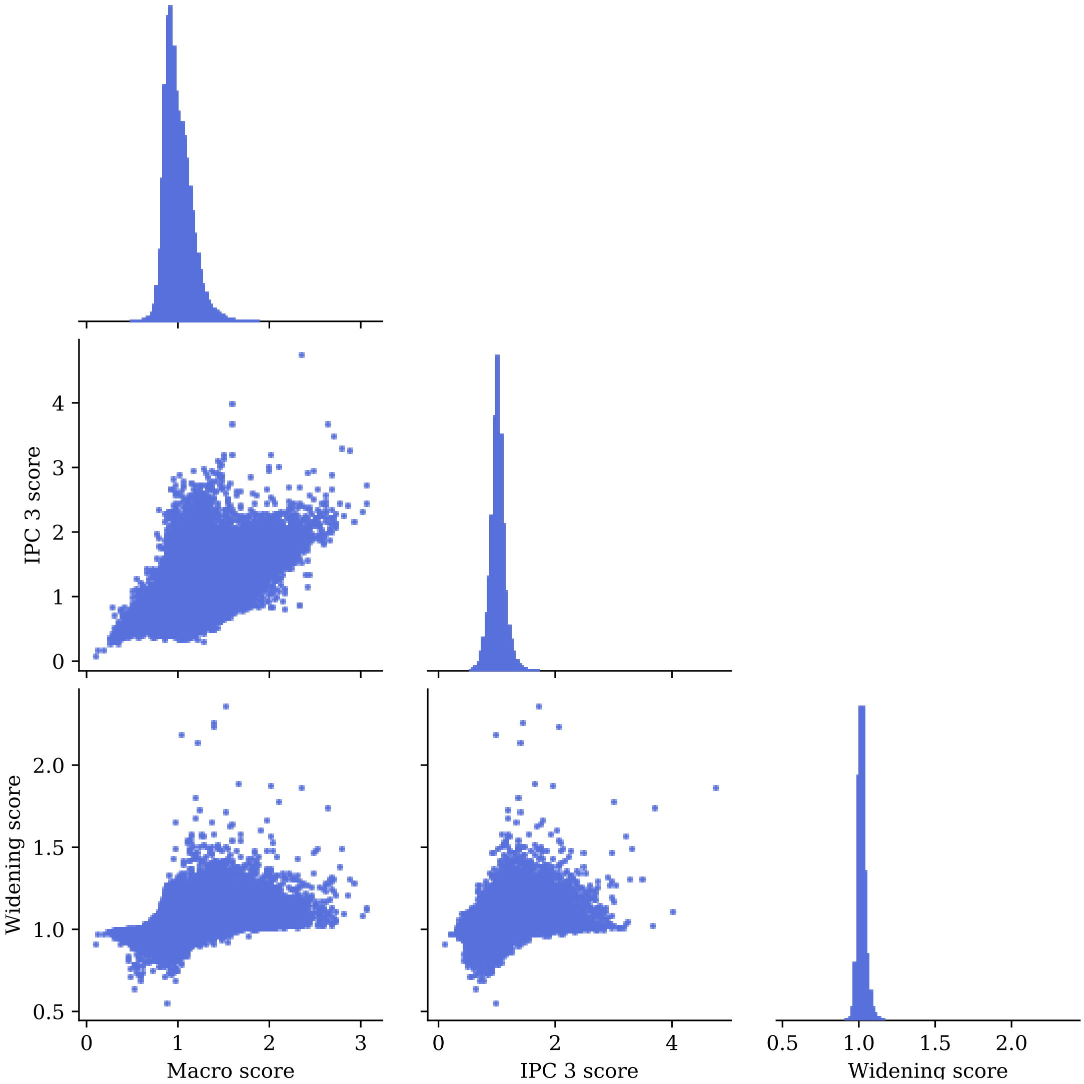
Note: the size of a word is proportional to its frequency. Words frequencies are obtained by averaging columns of the dtm over the subset of rows corresponding to the patents of interest: in panel (A), patents filed in 1976-1985 in IPC 3 C01; in panel (B), patents filed in 2006-2015 in IPC 3 C01.

FIGURE 2.B.5: BACKWARD VERSUS FORWARD SIMILARITIES


(a) Macro score
(b) IPC 3 score
(c) Widening score

Note: scatter plots of backward and forward similarities for each score type: macro, IPC3 and widening. Each dot is a patent.

FIGURE 2.B.6: CORRELOGRAM OF SCORES

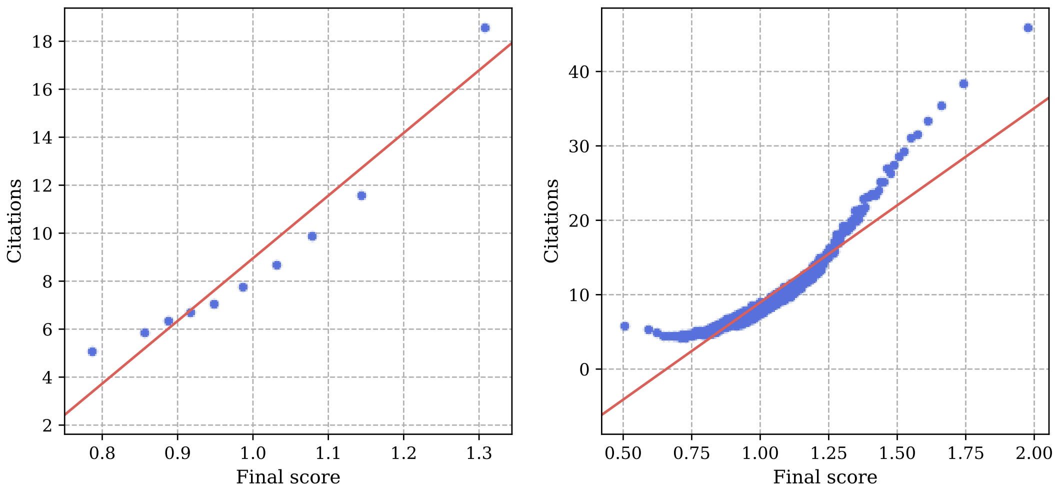


Note: the histograms of the scores are displayed on the diagonal, and scatter plots of scores against each other are displayed off the diagonal. Each dot is a patent.

Figure 2.B.7: Macro scores versus citations

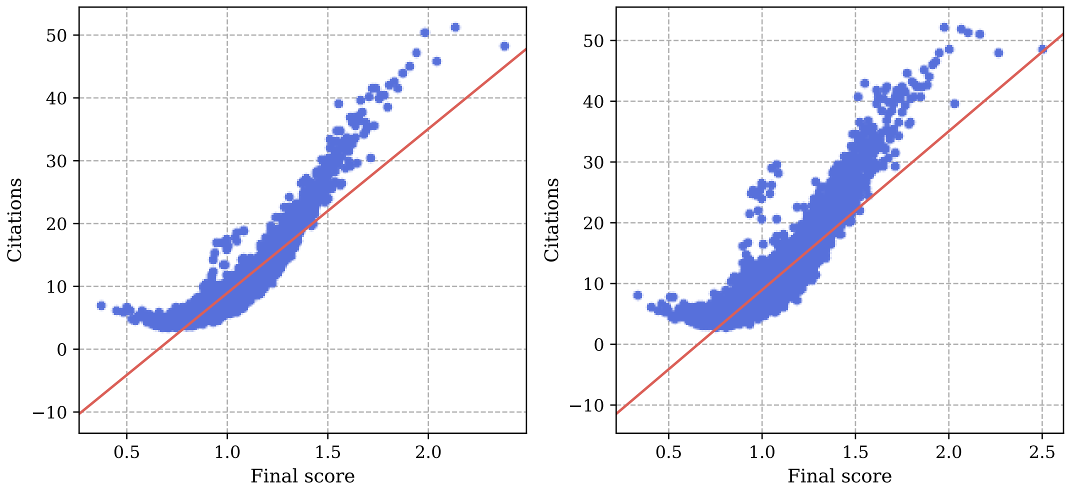
(a) Number of bins=10

(b) Number of bins=1000



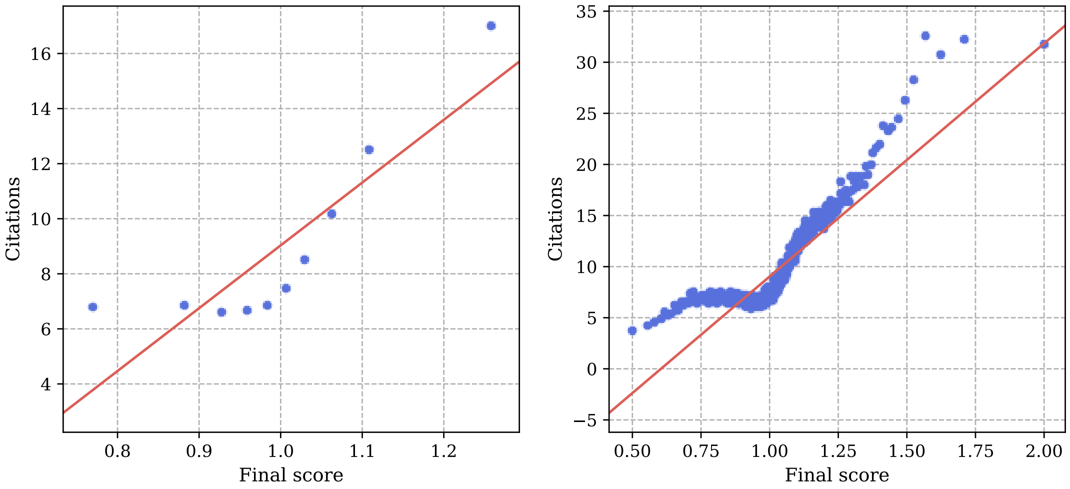
(c) Number of bins=10000

(c) Number of bins=20000



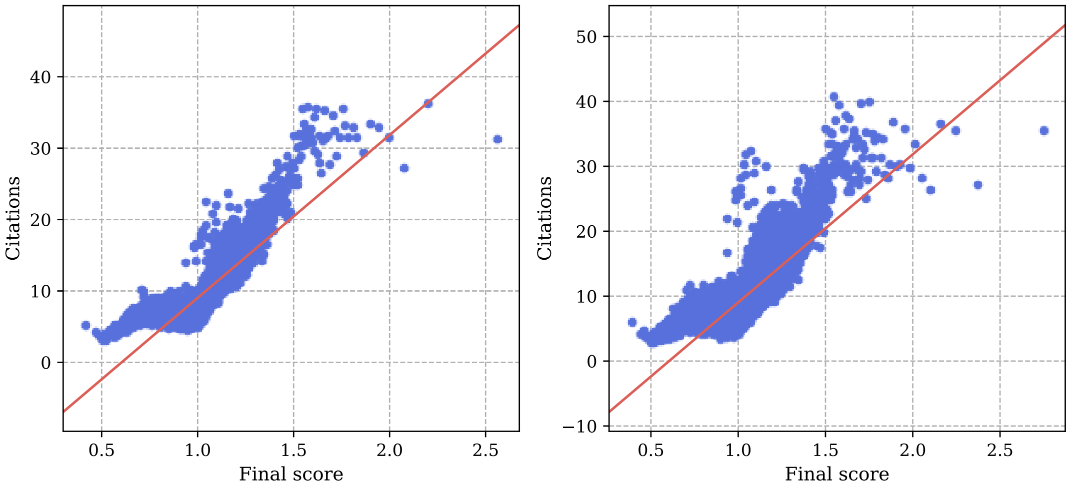
Note: binned scatter of scores against 10-year forward citations, with varying number of bins. Each dot is a patent.

FIGURE 2.B.8: IPC 3 SCORES VERSUS CITATIONS



(a) Number of bins=10

(b) Number of bins=1000

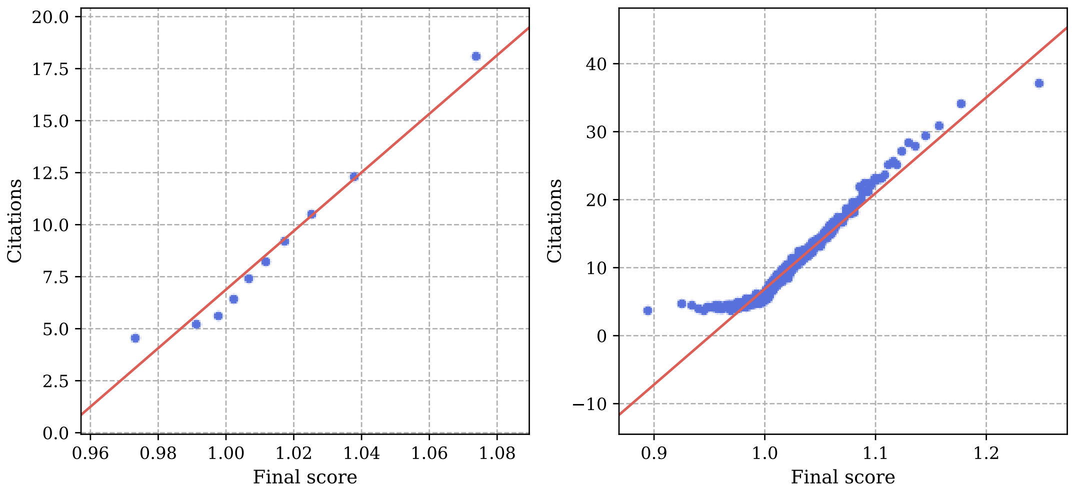


(c) Number of bins=10000

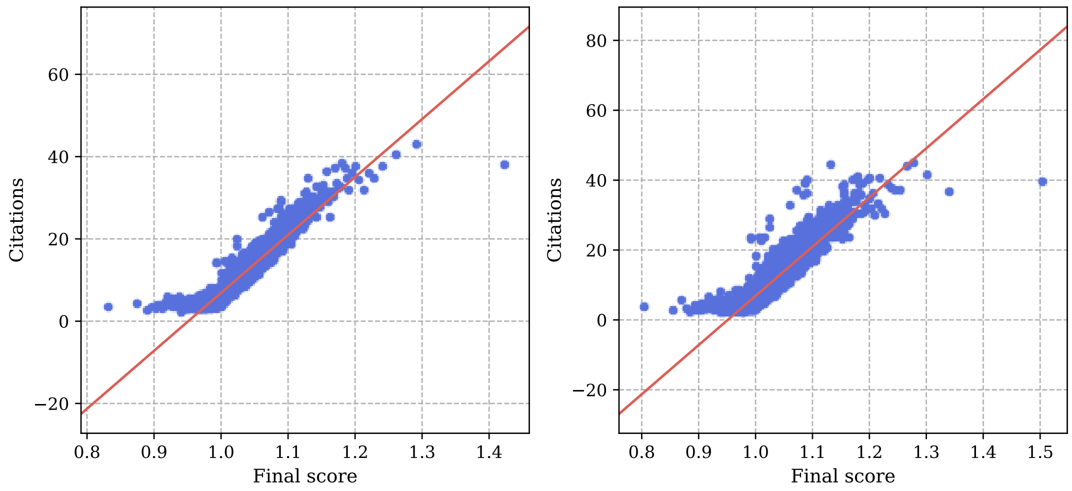
(c) Number of bins=20000

Note: binned scatter of scores against 10-year forward citations, with varying number of bins. Each dot is a patent.

FIGURE 2.B.9: WIDENING SCORES VERSUS CITATIONS


(a) Number of bins=10

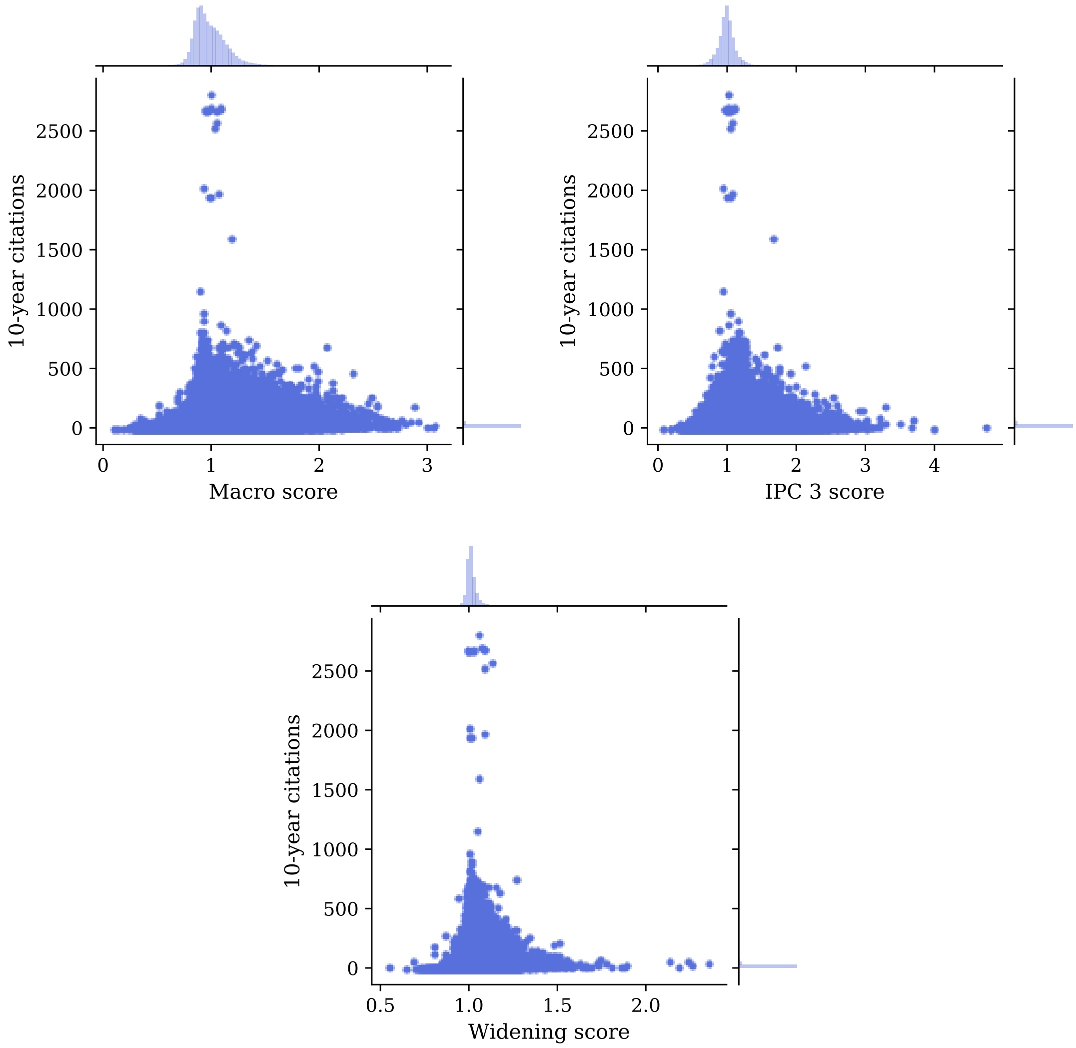
(b) Number of bins=1000


(c) Number of bins=10000

(c) Number of bins=20000

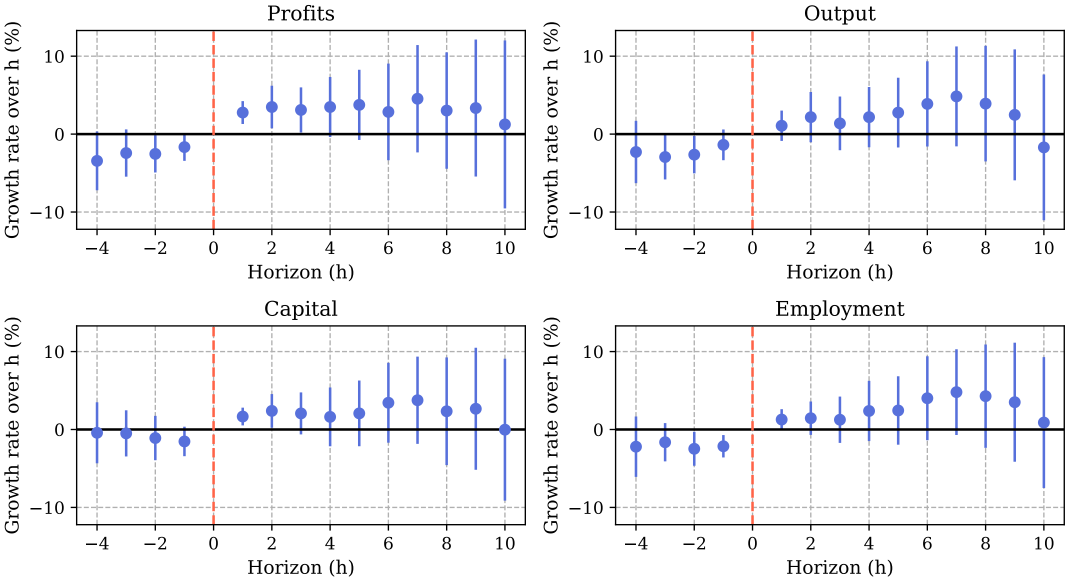
Note: binned scatter of scores against 10-year forward citations, with varying number of bins. Each dot is a patent.

FIGURE 2.B.10: SCATTER PLOTS OF SCORES AGAINST CITATIONS: RAW DATA



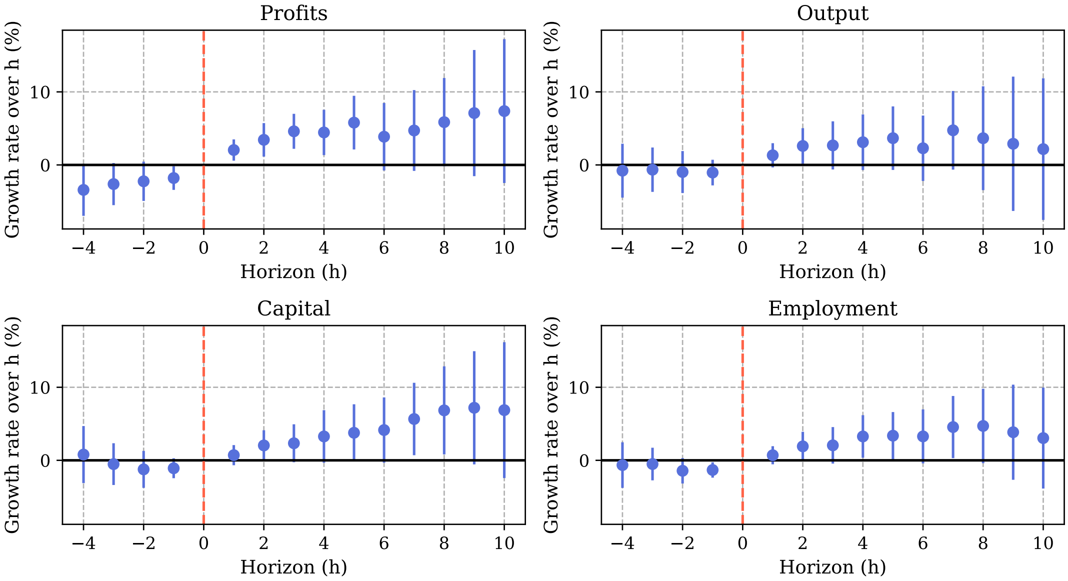
Note: scatter plots of 10-year citations against scores. The histograms of each variables are displayed on the outside of the boxes. Each dot is a patent.

FIGURE 2.B.11: TOP MACRO PATENTS AND FIRMS DYNAMICS: CUMULATIVE GROWTH RATES



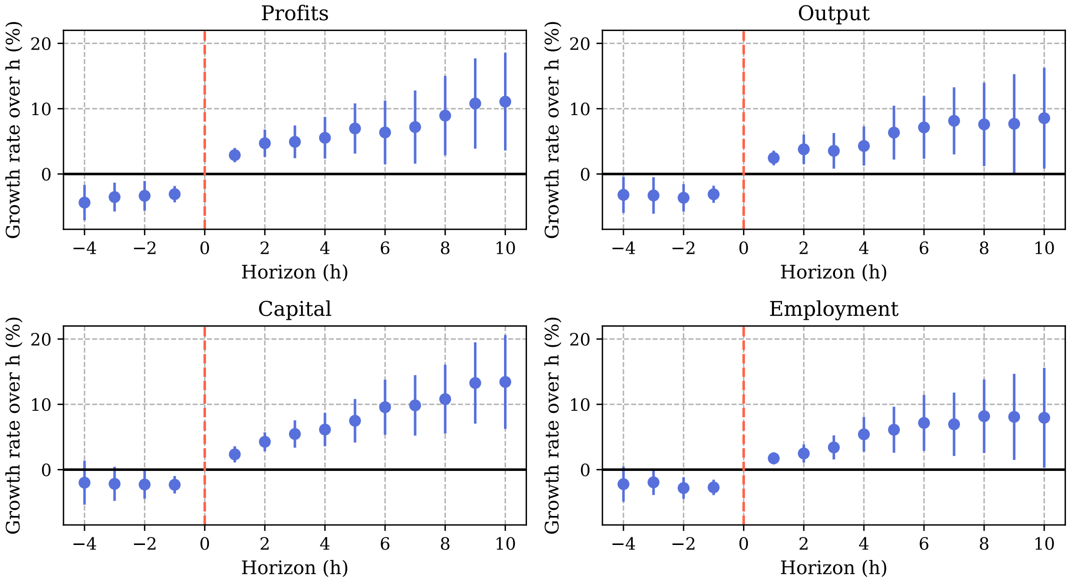
Note: estimates from equation (2.4) with alternative definition of growth rate using the macro score to qualify top patents. Dependent variable is $\log Y_{fi,t+h} - \log Y_{fi,t}$, i.e. the growth rate of the outcome value between time 0 and h. 95% confidence intervals are depicted. The coefficient is the growth rate between year $t+h$ and t using t as the base year, so for negative h a negative coefficient implies a positive growth rate between $t+h$ and t.

FIGURE 2.B.12: TOP IPC 3 PATENTS AND FIRMS DYNAMICS: CUMULATIVE GROWTH RATES



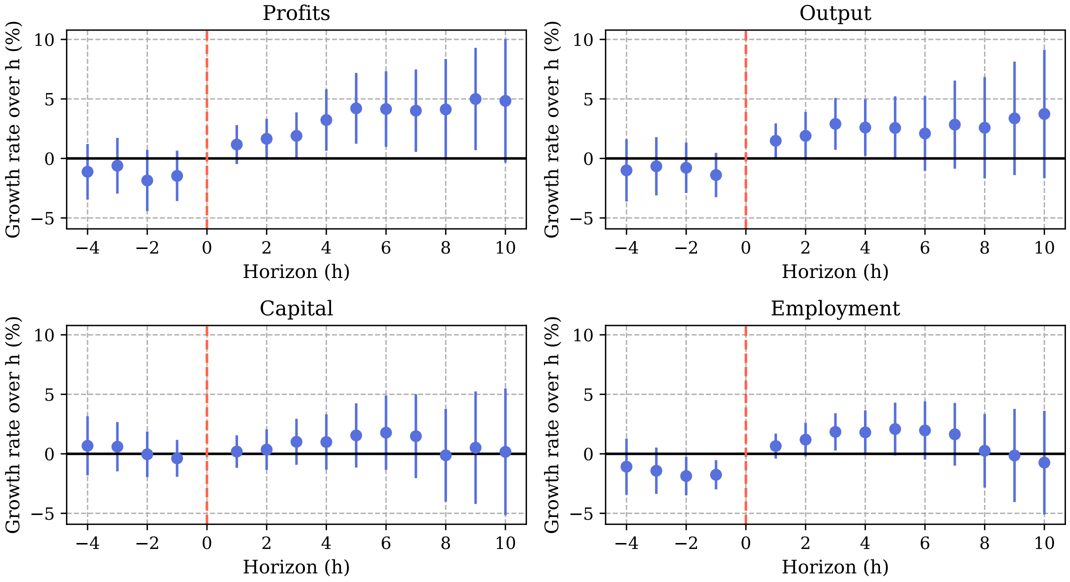
Note: estimates from equation (2.4) with alternative definition of growth rate using the IPC 3 score to qualify top patents. Dependent variable is $\log Y_{fi,t+h} - \log Y_{fi,t}$, i.e. the growth rate of the outcome value between time 0 and h. 95% confidence intervals are depicted. The coefficient is the growth rate between year $t+h$ and t using t as the base year, so for negative h a negative coefficient implies a positive growth rate between $t+h$ and t.

FIGURE 2.B.13: WIDENING PATENTS AND FIRMS DYNAMICS: CUMULATIVE GROWTH RATES



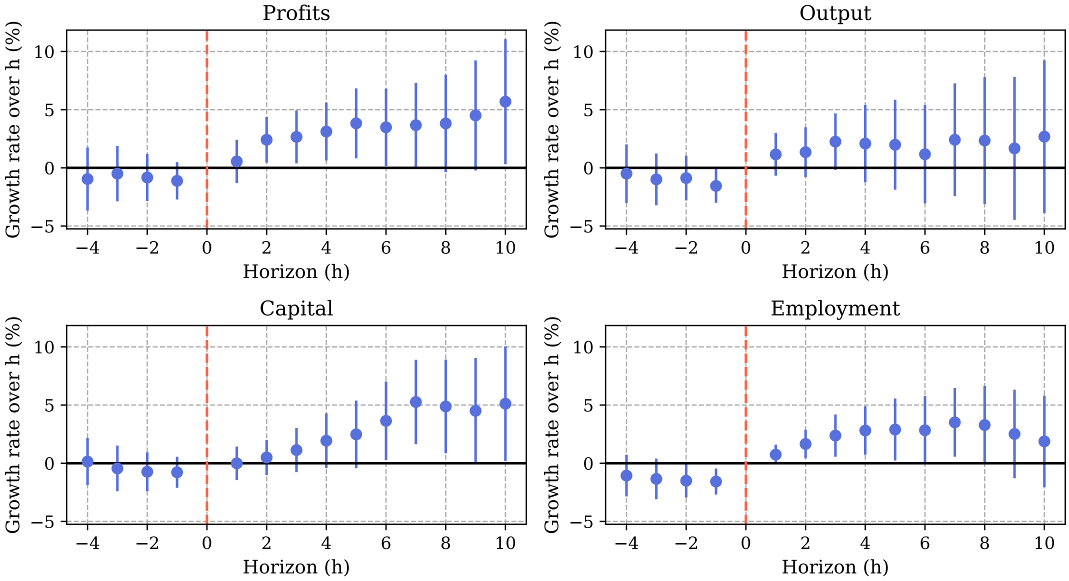
Note: estimates from equation (2.4) with alternative definition of growth rate using the widening score to qualify top patents. Dependent variable is $\log Y_{fi,t+h} - \log Y_{fi,t}$, i.e. the growth rate of the outcome value between time 0 and h. 95% confidence intervals are depicted. The coefficient is the growth rate between year $t+h$ and t using t as the base year, so for negative h a negative coefficient implies a positive growth rate between $t+h$ and t.

FIGURE 2.B.14: TOP MACRO PATENTS AND FIRMS DYNAMICS: WITHOUT IT



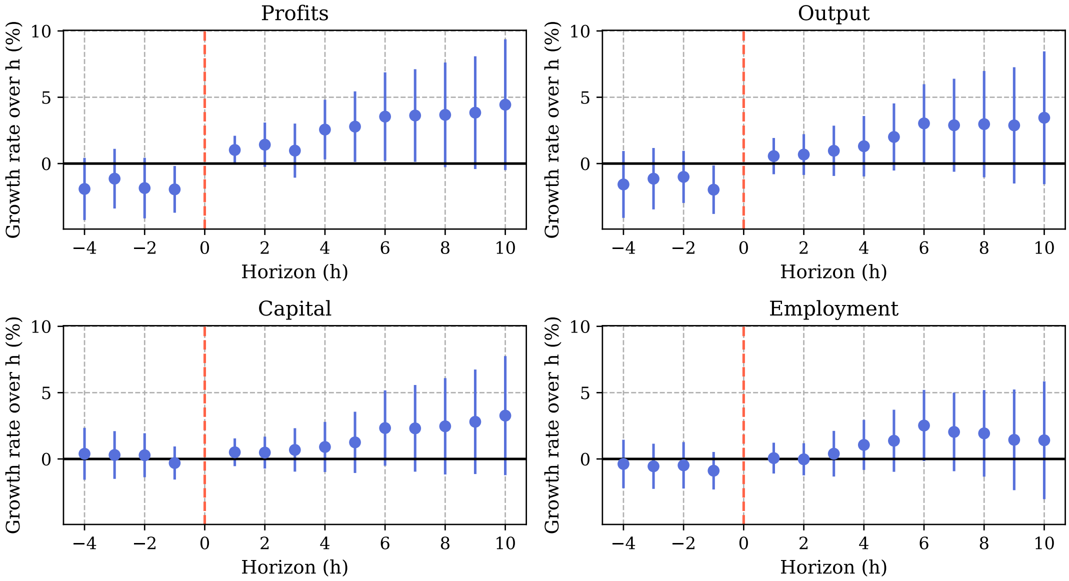
Note: estimates from equation (2.4) using the macro score to qualify top patents. 95% confidence intervals are depicted. The coefficient is the growth rate between year $t+h$ and t using t as the base year, so for negative h a negative coefficient implies a positive growth rate between $t+h$ and t. The sample excludes patents from IPC 1 codes G and H.

FIGURE 2.B.15: TOP IPC 3 PATENTS AND FIRMS DYNAMICS: WITHOUT IT



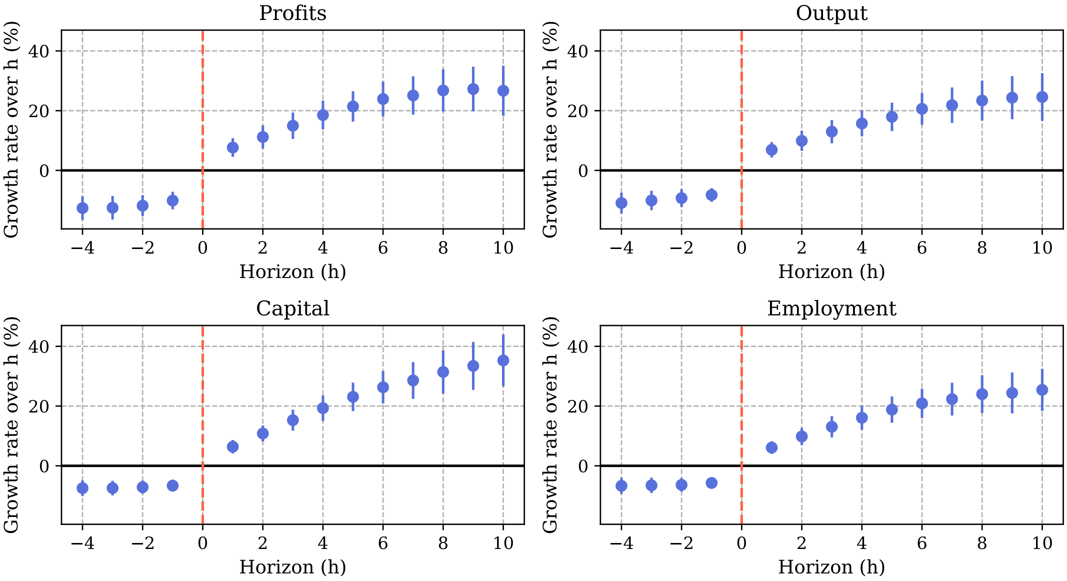
Note: estimates from equation (2.4) using the IPC 3 score to qualify top patents. 95% confidence intervals are depicted. The coefficient is the growth rate between year $t + h$ and t using t as the base year, so for negative h a negative coefficient implies a positive growth rate between $t + h$ and t. The sample excludes patents from IPC 1 codes G and H.

FIGURE 2.B.16: TOP WIDENING PATENTS AND FIRMS DYNAMICS: WITHOUT IT



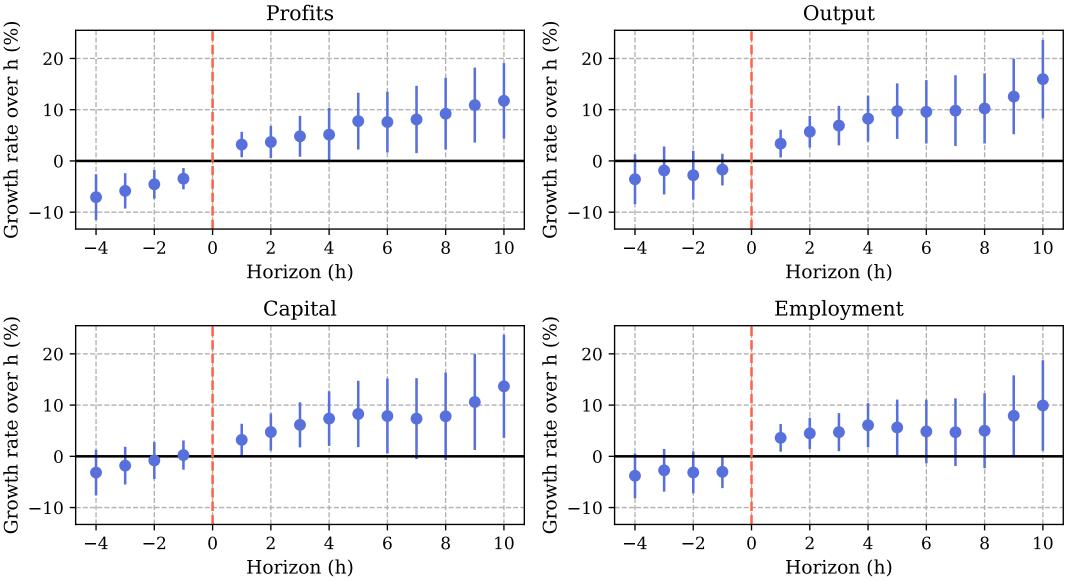
Note: estimates from equation (2.4) using the widening score to qualify top patents. 95% confidence intervals are depicted. The coefficient is the growth rate between year $t + h$ and t using t as the base year, so for negative h a negative coefficient implies a positive growth rate between $t + h$ and t. The sample excludes patents from IPC 1 codes G and H.

FIGURE 2.B.17: TOP PRIVATE VALUE PATENTS AND FIRMS DYNAMICS



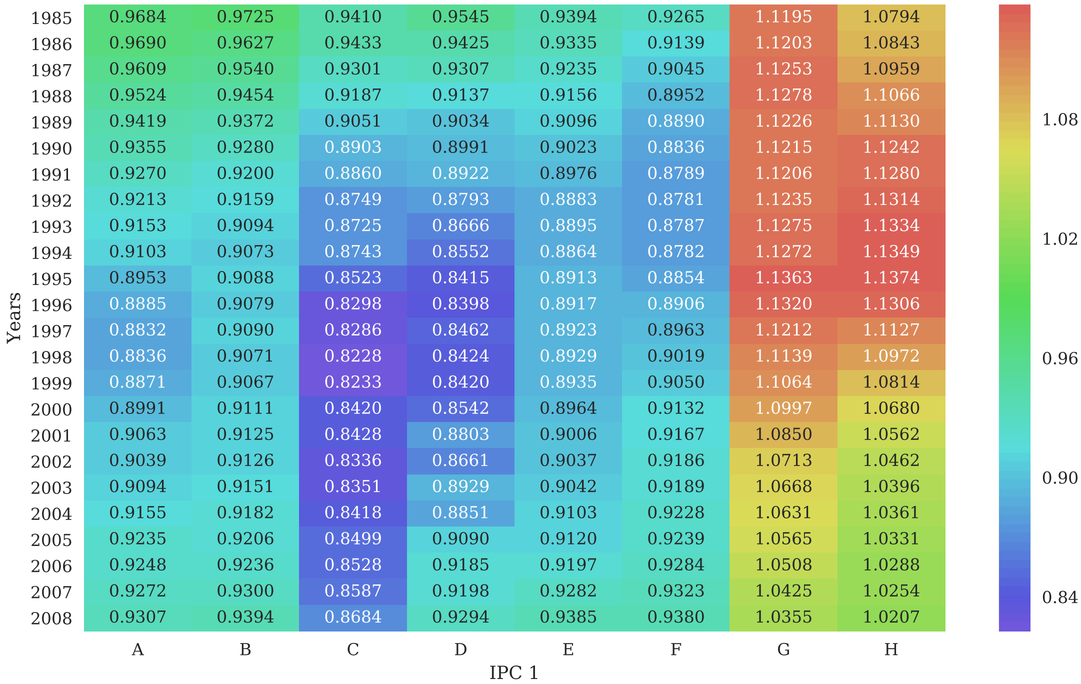
Note: estimates from equation (2.4) using private value from Stoffman et al. (2019) to qualify top patents. A top patent is one in the top 5% of the private value distribution (controlling for year fixed effects). 95% confidence intervals are depicted. The coefficient is the growth rate between year $t+h$ and t using t as the base year, so for negative h a negative coefficient implies a positive growth rate between $t+h$ and t.

FIGURE 2.B.18: TOP PATENTS IN TERMS OF CITATIONS AND FIRMS DYNAMICS



Note: estimates from equation (2.4) using citations at the 10-year horizon to qualify top patents. A top patent is one in the top 0.1% of the citations distribution (controlling for year fixed effects). 95% confidence intervals are depicted. The coefficient is the growth rate between year $t+h$ and t using t as the base year, so for negative h a negative coefficient implies a positive growth rate between $t+h$ and t.

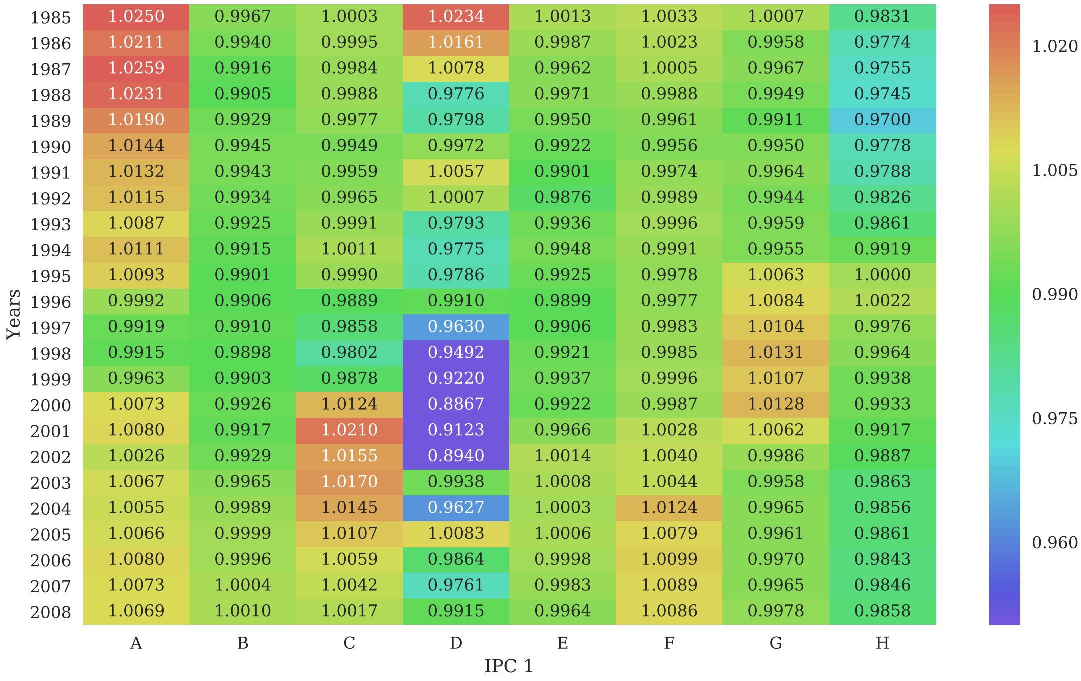
FIGURE 2.B.19: MACRO SCORES



Years	A	B	C	D	E	F	G	H
1985	0.9684	0.9725	0.9410	0.9545	0.9394	0.9265	1.1195	1.0794
1986	0.9690	0.9627	0.9433	0.9425	0.9335	0.9139	1.1203	1.0843
1987	0.9609	0.9540	0.9301	0.9307	0.9235	0.9045	1.1253	1.0959
1988	0.9524	0.9454	0.9187	0.9137	0.9156	0.8952	1.1278	1.1066
1989	0.9419	0.9372	0.9051	0.9034	0.9096	0.8890	1.1226	1.1130
1990	0.9355	0.9280	0.8903	0.8991	0.9023	0.8836	1.1215	1.1242
1991	0.9270	0.9200	0.8860	0.8922	0.8976	0.8789	1.1206	1.1280
1992	0.9213	0.9159	0.8749	0.8793	0.8883	0.8781	1.1235	1.1314
1993	0.9153	0.9094	0.8725	0.8666	0.8895	0.8787	1.1275	1.1334
1994	0.9103	0.9073	0.8743	0.8552	0.8864	0.8782	1.1272	1.1349
1995	0.8953	0.9088	0.8523	0.8415	0.8913	0.8854	1.1363	1.1374
1996	0.8885	0.9079	0.8298	0.8398	0.8917	0.8906	1.1320	1.1306
1997	0.8832	0.9090	0.8286	0.8462	0.8923	0.8963	1.1212	1.1127
1998	0.8836	0.9071	0.8228	0.8424	0.8929	0.9019	1.1139	1.0972
1999	0.8871	0.9067	0.8233	0.8420	0.8935	0.9050	1.1064	1.0814
2000	0.8991	0.9111	0.8420	0.8542	0.8964	0.9132	1.0997	1.0680
2001	0.9063	0.9125	0.8428	0.8803	0.9006	0.9167	1.0850	1.0562
2002	0.9039	0.9126	0.8336	0.8661	0.9037	0.9186	1.0713	1.0462
2003	0.9094	0.9151	0.8351	0.8929	0.9042	0.9189	1.0668	1.0396
2004	0.9155	0.9182	0.8418	0.8851	0.9103	0.9228	1.0631	1.0361
2005	0.9235	0.9206	0.8499	0.9090	0.9120	0.9239	1.0565	1.0331
2006	0.9248	0.9236	0.8528	0.9185	0.9197	0.9284	1.0508	1.0288
2007	0.9272	0.9300	0.8587	0.9198	0.9282	0.9323	1.0425	1.0254
2008	0.9307	0.9394	0.8684	0.9294	0.9385	0.9380	1.0355	1.0207

Note: each cell is the average macro score, by year and IPC section (IPC 1). Cells are colored such that higher values have hotter colors

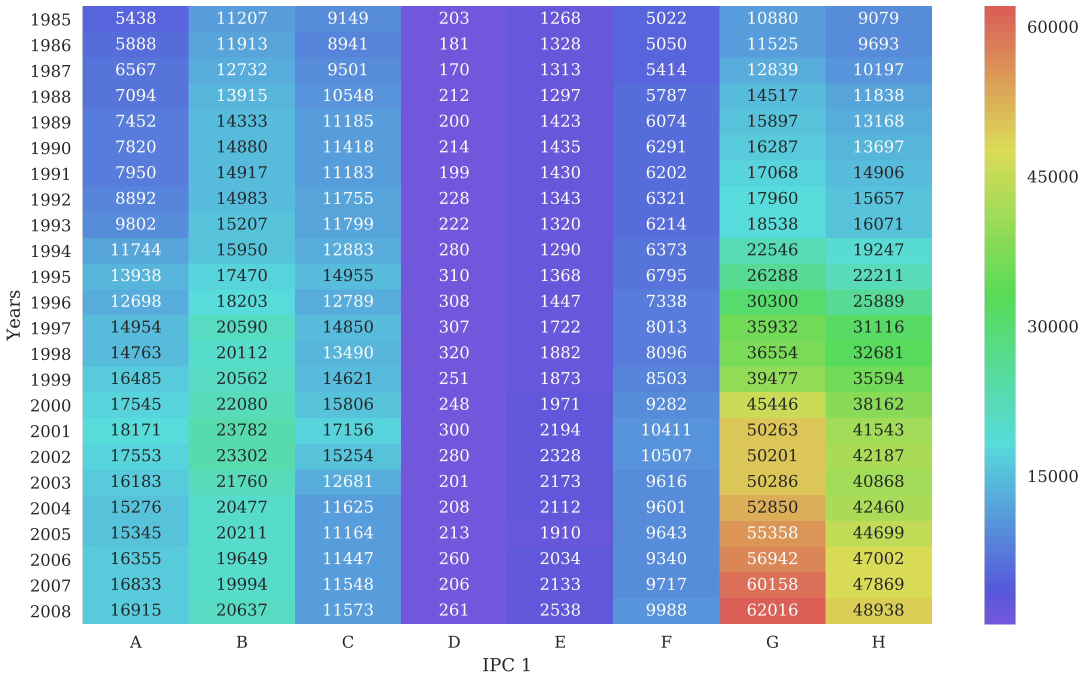
FIGURE 2.B.20: IPC 3 SCORES



Years	A	B	C	D	E	F	G	H
1985	1.0250	0.9967	1.0003	1.0234	1.0013	1.0033	1.0007	0.9831
1986	1.0211	0.9940	0.9995	1.0161	0.9987	1.0023	0.9958	0.9774
1987	1.0259	0.9916	0.9984	1.0078	0.9962	1.0005	0.9967	0.9755
1988	1.0231	0.9905	0.9988	0.9776	0.9971	0.9988	0.9949	0.9745
1989	1.0190	0.9929	0.9977	0.9798	0.9950	0.9961	0.9911	0.9700
1990	1.0144	0.9945	0.9949	0.9972	0.9922	0.9956	0.9950	0.9778
1991	1.0132	0.9943	0.9959	1.0057	0.9901	0.9974	0.9964	0.9788
1992	1.0115	0.9934	0.9965	1.0007	0.9876	0.9989	0.9944	0.9826
1993	1.0087	0.9925	0.9991	0.9793	0.9936	0.9996	0.9959	0.9861
1994	1.0111	0.9915	1.0011	0.9775	0.9948	0.9991	0.9955	0.9919
1995	1.0093	0.9901	0.9990	0.9786	0.9925	0.9978	1.0063	1.0000
1996	0.9992	0.9906	0.9889	0.9910	0.9899	0.9977	1.0084	1.0022
1997	0.9919	0.9910	0.9858	0.9630	0.9906	0.9983	1.0104	0.9976
1998	0.9915	0.9898	0.9802	0.9492	0.9921	0.9985	1.0131	0.9964
1999	0.9963	0.9903	0.9878	0.9220	0.9937	0.9996	1.0107	0.9938
2000	1.0073	0.9926	1.0124	0.8867	0.9922	0.9987	1.0128	0.9933
2001	1.0080	0.9917	1.0210	0.9123	0.9966	1.0028	1.0062	0.9917
2002	1.0026	0.9929	1.0155	0.8940	1.0014	1.0040	0.9986	0.9887
2003	1.0067	0.9965	1.0170	0.9938	1.0008	1.0044	0.9958	0.9863
2004	1.0055	0.9989	1.0145	0.9627	1.0003	1.0124	0.9965	0.9856
2005	1.0066	0.9999	1.0107	1.0083	1.0006	1.0079	0.9961	0.9861
2006	1.0080	0.9996	1.0059	0.9864	0.9998	1.0099	0.9970	0.9843
2007	1.0073	1.0004	1.0042	0.9761	0.9983	1.0089	0.9965	0.9846
2008	1.0069	1.0010	1.0017	0.9915	0.9964	1.0086	0.9978	0.9858

Note: each cell is the average IPC 3 score, by year and IPC section (IPC 1). Cells are colored such that higher values have hotter colors

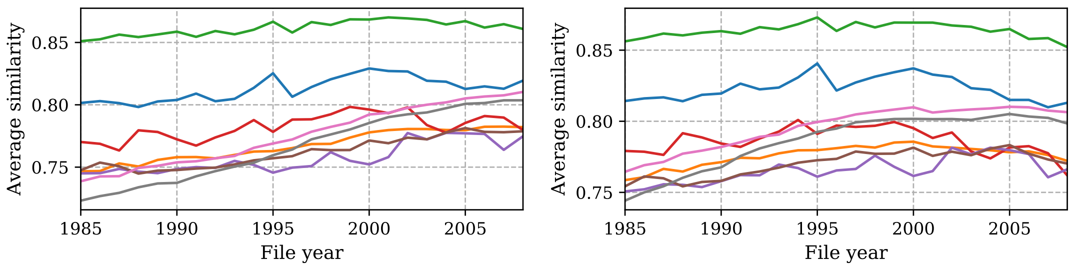
FIGURE 2.B.21: OBSERVATIONS (TOTAL: 2,745,260)



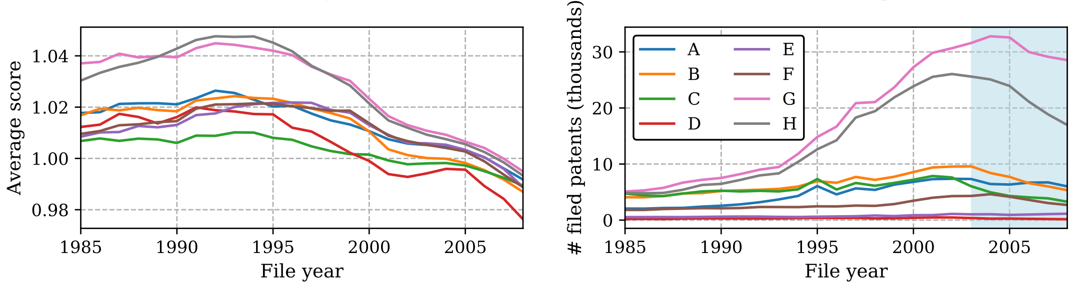
	A	B	C	D	E	F	G	H
1985	5438	11207	9149	203	1268	5022	10880	9079
1986	5888	11913	8941	181	1328	5050	11525	9693
1987	6567	12732	9501	170	1313	5414	12839	10197
1988	7094	13915	10548	212	1297	5787	14517	11838
1989	7452	14333	11185	200	1423	6074	15897	13168
1990	7820	14880	11418	214	1435	6291	16287	13697
1991	7950	14917	11183	199	1430	6202	17068	14906
1992	8892	14983	11755	228	1343	6321	17960	15657
1993	9802	15207	11799	222	1320	6214	18538	16071
1994	11744	15950	12883	280	1290	6373	22546	19247
1995	13938	17470	14955	310	1368	6795	26288	22211
1996	12698	18203	12789	308	1447	7338	30300	25889
1997	14954	20590	14850	307	1722	8013	35932	31116
1998	14763	20112	13490	320	1882	8096	36554	32681
1999	16485	20562	14621	251	1873	8503	39477	35594
2000	17545	22080	15806	248	1971	9282	45446	38162
2001	18171	23782	17156	300	2194	10411	50263	41543
2002	17553	23302	15254	280	2328	10507	50201	42187
2003	16183	21760	12681	201	2173	9616	50286	40868
2004	15276	20477	11625	208	2112	9601	52850	42460
2005	15345	20211	11164	213	1910	9643	55358	44699
2006	16355	19649	11447	260	2034	9340	56942	47002
2007	16833	19994	11548	206	2133	9717	60158	47869
2008	16915	20637	11573	261	2538	9988	62016	48938

Note: each cell is the total number of patents filed, by year and IPC section (IPC 1). Cells are colored such that higher values have hotter colors

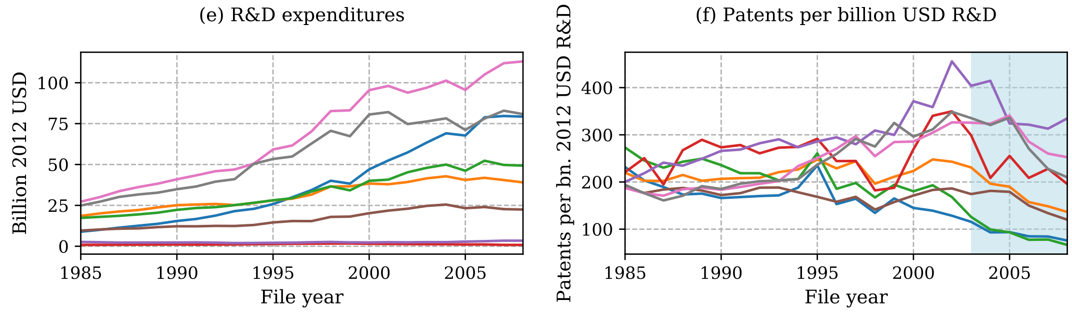
FIGURE 2.B.22: TRENDS BY IPC 1 — PATENTS LINKED TO FIRMS


(a) Backward neighbourhood similarity

(b) Forward neighbourhood similarity


(c) Widening score

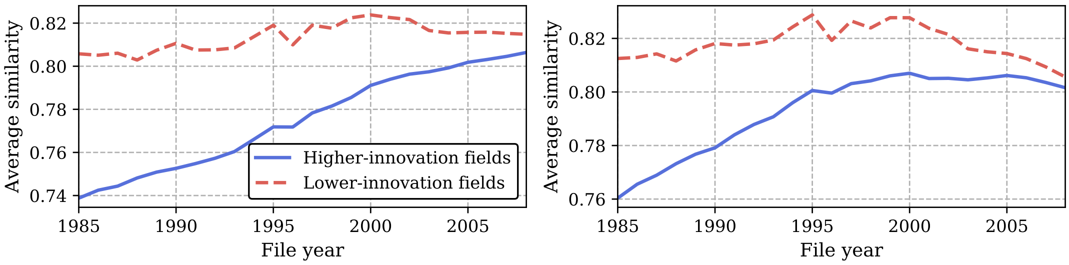
(d) Number of patents


(e) R&D expenditures

(f) Patents per billion USD R&D

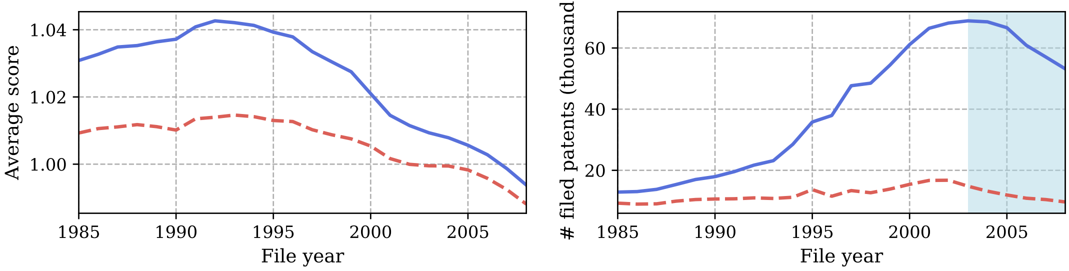
Note: panel (a): average backward neighbourhood similarity by IPC 1 over time; panel (b): average forward neighbourhood similarity by IPC 1 over time; panel (c): average widening score by IPC 1 over time; panel (d): number of patents filed by IPC 1 over time; panel (e): firms' R&D expenditure in each IPC 1 field; panel (f): patents per dollar of R&D expenditure (contemporaneous). Sample: patents linked to Compustat firms. The shaded area corresponds to years in which a decrease in the number of patents filed by Compustat firms is observed, yet it is unclear whether this decrease is due to fewer patent-firm matches or to an actual decrease in filing. R&D expenditures at the firm level are allocated to IPC 1 fields based on the fraction of patents filed in each field in the 10 following years.

FIGURE 2.B.23: TRENDS BY IPC 3 GROUP — PATENTS LINKED TO FIRMS



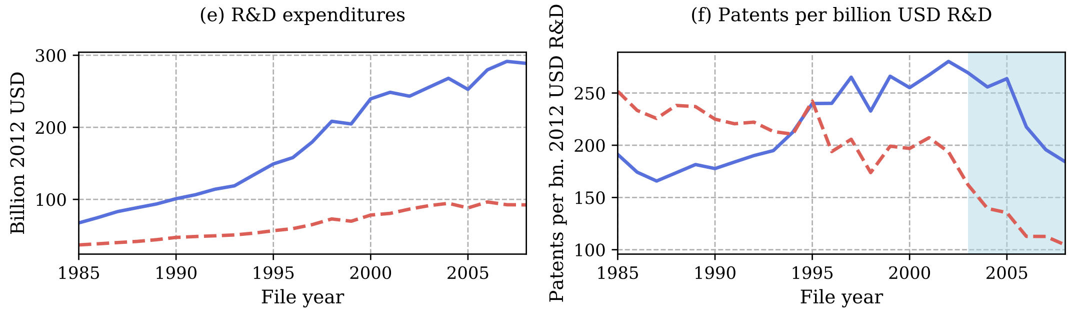
(a) Backward neighbourhood similarity

(b) Forward neighbourhood similarity



(c) Widening score

(d) Number of patents



(e) R&D expenditures

(f) Patents per billion USD R&D

Note: panel (a): average backward neighbourhood similarity by IPC 3 group over time; panel (b): average forward neighbourhood similarity by IPC group over time; panel (c): average widening score by IPC 3 group over time; panel (d): number of patents filed by IPC 3 group over time; panel (e): firms' R&D expenditure in each IPC group; panel (f): patents per dollar of R&D expenditure (contemporaneous). Sample: patents linked to Compustat firms. The shaded area corresponds to years in which a decrease in the number of patents filed by Compustat firms is observed, yet it is unclear whether this decrease is due to fewer patent-firm matches or to an actual decrease in filing R&D expenditures at the firm level are allocated to IPC 1 fields based on the fraction of patents filed in each field in the 10 following years.

Chapter 3

Cross-border Value Added Tax Fraud In The European Union

3.1 Introduction

Value Added Tax (VAT) is an important source of revenues in the European Union (EU). In 2018, VAT revenues accounted for 17.4% of total tax revenues of Member States and for more than half of revenues from taxes on production and imports.[1] Every year, Member States experience large losses in VAT revenues to fraud and non-compliance. Official statistics by the European Commission indicate that EUR 137.5 billion in potential VAT revenues were lost in 2017, about 11% of the theoretical tax liability (Poniatowski et al., 2019).

This paper studies the effects of a reform of the EU VAT rules that removes the possibility for firms to engage in a set of well-known fraudulent schemes. These fraud strategies exploit a weakness in the way within-EU cross-border transactions are treated for VAT purposes, which allows fraudsters to collect VAT without rightfully remitting it to the tax authorities. Crucially, international trade flows are misreported as a direct consequence of these fraudulent manoeuvres. This paper proposes a quantification of the amount of fraud the reform eliminates by tracking discrepancies between trade statistics reported by importing and exporting countries around the reform time. This is an important question in the EU policy circles (European Commission, 2018) and the present exercise is indirectly informative of the levels of VAT fraud prevailing in the EU.[2]

---

[1] Source: Eurostat, database *gov_10a_taxag*.

[2] European Commission (2018, p. 5) relates that "A short survey was sent to all Member States to

The fraud schemes studied in this paper exploit two features of the VAT system (de la Feria, 2019). First, cross-border transactions are subject to the *Reverse Charge Mechanism* which shifts the liability of the tax from the seller onto the buyer. In other words, whereas the seller normally charges VAT on its sale and remits it to the taxman, it is the responsibility of the buyer to pay VAT on its purchase to its own national tax authorities under the Reverse Charge Mechanism — exports are zero-rated and imports are subject to domestic VAT. This is consistent with the *destination principle*: goods are taxed in the country where they are sold.[3] Second, the absence of customs between EU countries removes the ability to collect VAT at the border, shifting instead the duty to report to the buyer and resulting in a time lag between the time at which the transaction takes place and that of taxation.

I consider three types of fraud enabled by these features. First, Missing Trader Intra-Community (MTIC) fraud, in which a firm imports goods from another Member State and sells it domestically, collecting VAT on the sale. The firm then disappears without remitting the VAT from its sale and without having paid VAT on its import.[4] This scheme is mostly run by criminal organisations and is believed to cause governments large revenue losses. Second, domestic sales being falsely reported as exports, which are VAT-free. The fraudster pockets the VAT on its domestic sales as the authorities do not expect to collect any VAT on exports. Third, imports reported as domestic purchases: since VAT is rebated on purchases (irrespective of the origin) but it has not been paid on imports at the time of transaction, disguising imports as domestic purchases allows the firm to benefit from a rebate for VAT it has never paid. Each of these fraud schemes results in misreported trade flows.[5]

The reform allows Member States to impose the Reverse Charge Mechanism on *domestic transactions* of selected goods, rendering all fraud schemes above ineffective. MTIC fraud is removed as firms cannot charge VAT to their customers anymore and therefore become unable to run away with any taxed amounts. Disguising domestic sales as exports is not profitable anymore since VAT cannot be charged to customers

---

enquire about their experiences with estimating VAT fraud and/or Missing Trader Intra-Community (MTIC) fraud. The results showed that there is little experience in estimating the size of these types of fraud."

[3] The destination principles applies to most international transactions in most countries.

[4] The key difference between MTIC fraud and misreporting sales of goods sourced domestically is the fact that firms dodge the payment of VAT on imports whereas it is much harder to do so for domestic purchases. In MTIC schemes, the fraudsters gain VAT on whole the sale value and not only on the value added. More details are given in section 3.3.2.

[5] Formally, imports and exports within the EU are called *arrivals* and *dispatches*, respectively. I use imports and exports throughout the paper for simplicity.

irrespective of their country of establishment. Last, the gain from misrepresenting imports as domestic purchases vanishes since purchasers are also liable for VAT on their purchases made domestically.

I collected data on 54 instances of implementation of the *Domestic Reverse Charge Mechanism* (henceforth abbreviated as DRCM) spanning different goods across 24 Member States over years 2004-2019. To measure the effect of the reform on fraud, I exploit an empirical strategy originally used in the literature on tariff evasion in international trade (Fisman and Wei, 2004). In trade statistics, each trade flow is reported twice: once by the exporter as an export, and once by the importer as an import. The difference between these reported figures (reported exports minus reported imports) is defined as a *reporting gap*. When the reporting incentive changes for one side — e.g. because the level of tariff changes, inducing importers to lower their reported imports to pay less tariff — but remains constant for the trading partner, the reporting gap changes as a result. Identification of fraud therefore relies on the reaction of the reporting gaps — computed using aggregate trade flows between countries at the 8-digit product level — to the implementation of the DRCM, which changes the reporting incentives of the fraudsters located in the country of reform. I use a difference-in-difference approach comparing the reporting gaps associated with the goods to which the DRCM applies (treated) to those to which it does not (control) within a narrowly defined time window around the reform date.

This exercise presents challenges that I discuss at length in the text. Perhaps the most important one is the fact that goods to which the DRCM is applied are selected based on suspicion of high levels of fraud, i.e. the assignment is not random. In that sense, the inferred level of fraud must be understood as that prevailing for these goods at the time of reform, and cannot be taken as a precise estimate of the importance of VAT fraud in the EU.

I find that in the months following the reform, the reporting gaps associated with treated goods *imported* by the reform country decrease by around 3% on average relative to goods not subject to the DRCM. In other words, reported imports increase relative to reported exports. This is consistent with the removal of MTIC fraud and of the fraudulent disguise of imports as domestic purchases. The reporting gaps associated with treated goods *exported* by the reform country are not impacted at the time of reform. This suggests that fraud involving mis-classification of domestic sales as export is quantitatively less important. This is consistent with the view that MTIC-type schemes are the most important type of cross-border VAT fraud (Lamensch and Ceci, 2018).

There is considerable heterogeneity in the effects of the reform across the 54 episodes. In some instances, the reporting gaps associated with imported goods that are treated decrease by as much as 63%. In other rarer cases, the reporting gaps significantly *increase* in contradiction to all predictions. The inferred magnitudes of fraud prevailing pre-reform (as measured by the change in the reporting gaps) appear to be larger in countries with higher values of the VAT rate, corruption levels and VAT non-compliance, as well as in countries with less efficient customs. I do not find evidence of fraud shifting to other goods that are not subject to the DRCM.

I compute pre-reform fraud amounts based on the reaction of the reporting gaps at the time of reform and trade volumes of the treated products. I find that these figures are very sensitive to whether I use the average treatment effect across events or event-specific estimates. Summing across all events, estimated fraud in the year leading to the reform ranges from EUR 610 million to EUR 1.6 billion. Irrespective of the choice of parameters underlying these calculations, fraud removed by the reform is tiny compared to VAT revenues (0.06%-0.2% on average). This is surprisingly low in light of all the efforts undertaken by the EU to fight this type of VAT fraud.

Relation to the literature This paper contributes to the literature concerned with estimating VAT fraud, focusing on specific schemes that involve misreporting of cross-border transactions. Keen and Smith (2006) provide a review of known VAT schemes. Gradeva (2014) shows a correlation between VAT rate in the importing country and reporting gaps in the EU which she interprets as evidence of MTIC fraud, as the gains from fraud are proportional to the VAT rate. Benzarti and Tazhitdinova (2019) suggest that trade flows (i.e. real activity) do not appear to react to changes in VAT rates in the EU, concluding that VAT appear not to distort trade. Based on the reporting gap method also used in this paper, Ferrantino et al. (2012) show that in the case of China — which does not systematically fully rebates VAT on exports, unlike in the EU — traders under-report their exports to evade VAT.

Besides fraud involving cross-border transactions, VAT fraud attracted attention as this tax is believed to be self-enforcing in nature since buyers have an incentive to ask for receipts and report their purchases in order to get VAT refunds, which entails a paper trail. Pomeranz (2015) demonstrates that this paper trail has considerable enforcement power when combined with deterrence measures that incentivise firms to report accurately — e.g. a threat of audit. However, the last transaction to the final consumer is not subject to this self-enforcing mechanism. Naritomi (2019) studies the effects of rewarding consumers for reporting their purchases and finds that firms selling to final consumers increase their reported sales significantly in response. In the EU, Hopland and Ullmann (2019) finds that restaurants in Germany evade VAT by mis-classifying dine-in meals as take-away, the former being subject to the standard rate whereas the latter benefits from a reduced rate.

Zídková (2019) calculates the impact of one instance of the application of the DRCM on the trade balance between Czech Republic and the rest of the EU and interprets its decrease as evidence of MTIC fraud — assuming that MTIC-enabled trade shifts to other goods or countries upon implementation of the DRCM. The approach I take does not rely on trade balance data, but on the difference between reported imports and exports induced by fraud. The two papers closest to the present study are Stiller and Heinemann (2019) (in German) and CASE (2015) (in Polish). They study how the reporting gaps react to the DRCM in specific instances of the reform in Germany, Austria and Poland. My work differs from theirs in several ways. First, I use data at the 8-digit level of the Combined Nomenclature without aggregating trade flows, which allows more precision in both definition of the treated goods and in estimation. Second, I study all instances (to the best of my knowledge) in which the DRCM was applied.

This paper also build on studies that use reporting gaps to estimate tariff evasion. The idea to compare values of the same trade flow reported independently by the importing and exporting countries originates in the seminal work by Fisman and Wei (2004). Using data on trade between China and Hong Kong, the authors found evidence of tariff evasion by investigating the effect of changes in tariff rates on the reporting gaps in a cross-section of products. This approach was extensively used subsequently (see amongst others Mishra et al., 2008; Javorcik and Narciso, 2008; Stoyanov, 2012; Javorcik and Narciso, 2017; Kellenberg and Levinson, 2019; Levin and Widell, 2014; Farhad et al., 2018; Demir and Javorcik, 2020). In Bussy (2020), I use it to investigate corporate income tax evasion.

Structure of the paper Section 3.2 introduces the VAT rules in the EU. Fraud schemes involving cross-border transactions and their relation to the Domestic Reverse Charge Mechanism are explained in section 3.3. Section 3.4 lays out the theoretical link between the trade reporting gaps and cross-border VAT fraud. The data and the empirical strategy are introduced in sections 3.5 and 3.6, respectively. Section 3.7 details the results and section 3.8 concludes. Additional tables and figures, respectively referenced with prefixes 3.A and 3.B, can be found in appendices of the same name.

3.2 The Value Added Tax rules in the EU

In this section, I describe the VAT rules of the EU, and its details related to intra-community trade, i.e. trade taking place between two Member States. A detailed description of the VAT system in the EU is provided by Baldwin (2007). The VAT legislation is stipulated in Council Directive 2006/112/EC (VAT Directive henceforth) — and 77/388/EEC before 2006.

European Value Added Tax VAT is a type of consumption tax whose collection takes place incrementally along the production stages. VAT generally applies to all transactions in an economy: sellers charge VAT on their sales to buyers at every stage of production. The seller can deduct VAT from its own business purchases, so that the total VAT for which the seller is liable over a given period of time amounts to the VAT charged to its consumer, minus the VAT paid on its own purchases. Since business expenses are VAT-deductible, the tax base is aggregate consumption (Baldwin, 2007). Importantly, firms along the production chain are liable for VAT on their value added, but charge it to their customers. Ignoring the effects VAT can have on prices and the administrative costs of accounting for it, a firm that can fully deduct VAT on its inputs should be indifferent between living in an economy where VAT is in place, versus one where it is not. In practice, almost all purchases can be accounted as business purchases except expenses on luxuries, amusements or entertainment (art. 176 VAT Directive).

One attractive feature of the VAT is the apparent ease of its implementation. The tax authority does not chase after firms to collect the tax as buyers have an incentive to ensure that the transaction is reported accurately, since VAT on their purchases can be deducted from their final VAT liability. The effort in assessing the tax base is delegated from the tax authority to the buying firms. In practice, businesses are required to file their VAT returns several times a year, stating their sales and purchases over the tax period on the basis of which their VAT balance is determined.[6]

There may be several VAT rates in each country, and tax rates are currently not harmonized across countries. In addition to the standard rate — which applies to most products — each country can have up to two reduced rates that typically apply to products of first necessity, such a foodstuff or medical equipment. At the time of writing, the European Commission set a minimum standard rate of 15% and a minimum lowest reduced rate of 5%, limits above which EU members are free to set their national rates.

VAT on cross-border transactions It is important to distinguish between business-to-business (B2B) and business-to-consumer (B2C) transactions in intra-community cross-border trade. To qualify as B2B, the buying firm must be registered for VAT in the destination country. B2B transactions account for most of trade in terms of volumes and are subject to different rules than domestic transactions. Whereas domestic transactions are all subject to the VAT and each seller charges VAT on its sales, exports to a business located (and registered for VAT) in an another EU country are zero-rated (i.e. the exporter does not charge VAT).[7] The importer is liable for VAT on its imports from another member state. This follows the so-called *destination principle*, whereby the VAT is collected and kept by the country in which the goods are being sold — i.e. where the goods are located after the sale. In other words, there is no VAT on exports, but VAT is levied on imports. Because there are no customs, it is the responsibility of the importer to declare the transaction in its VAT return and pay the tax in the destination country at the rate applicable in that country. The shifting of the liability of VAT from the seller to the buyer is called *Reverse Charge Mechanism* and it applies to all within-EU international B2B transactions. The chain whereby VAT is systematically charged on sales and remitted to the authorities by the seller is broken, potentially leaving more scope for a variety of fraudulent schemes that I describe in section 3.3.

---

[6] The frequency at which VAT returns must be filed varies across countries. The EU VAT Directive allows tax periods of one, two or three months. For instance, the tax period is 3 months in the UK.

[7] Sales to consumers that do not have a valid VAT registration number in the destination country qualify as B2C. A firm selling to non-taxable consumers in another member state must register for VAT in the destination country and charge VAT at the applicable national rate, unless its total sales fall below a certain threshold (in which case it charges VAT at the rate applicable in its own country). This threshold is called a *threshold for distance selling* and its value vary across countries. The current thresholds can be found here (last accessed 19/05/2020).

3.3 VAT fraud in the European Union

VAT Fraud in the European Union is a serious issue that results in large revenue losses for governments. Fraud is difficult to measure due to its illegal nature and the most common measure of losses due to non-compliance is the *VAT Gap*, defined as the difference between the theoretical and the actual VAT liabilities.[8] It is available for all Member States, and constitutes a guide to policy when it comes to EU legislation. In 2017, the VAT Gap in the EU-28 amounted to EUR 137 billion or 11.2% of the theoretical tax liability — i.e. the tax liability according to the tax law (Poniatowski et al., 2019). Whilst the VAT Gap is small in some countries (e.g. ca. 0.6% in Cyprus and Luxembourg), it is large in others and exceeds 30% in Romania and Greece. However, fraud is only a fraction of the VAT Gap and its size is subject to intense debate, with estimates ranging from EUR 20 to EUR 100 billion annually (European Commission, 2018). Other factors such as mistakes, bankruptcies and insolvencies preventing firms from paying VAT are also contained in the VAT gap.

3.3.1 Types of VAT fraud

VAT fraud can take many forms. Keen and Smith (2006) and (European Commission, 2018) provide a comprehensive list of the known types of fraud on the VAT system. The rallying feature of these schemes is that firms collect VAT and do not remit it fully, or claim credit to which they are not entitled. One can generally divide fraud into two categories. On the hand, classical fraud schemes that include under-reporting sales, false claims for credit or refund, or credit claimed for VAT on non-business expenses (for which refund is not available). Any firm can engage in that type of fraud in principle. On the other hand, organized VAT fraud run by criminal organizations involving cross border transactions, that exploit weaknesses in the EU VAT system and that are described in the next section.

3.3.2 The Reverse Charge Mechanism on cross-border transactions as an enabler of fraud

The Reverse Charge Mechanism transfers the liability of the tax from the seller onto the buyer. In practice, this means that for a given transaction the seller zero-rates the

---

[8] For a critique of these methods, see Keen and Slemrod (2017); Gemmell and Hasseldine (2014).

sale (i.e. does not charge VAT) and the buyer owes VAT to the tax authorities. Since firms can claim refunds on VAT paid on purchases, the buyer does not pay any VAT in practice: the VAT owed is claimed back on the same form. The Reverse Charge Mechanism applies to transactions between Member States since the inception of the Single Market.[9] The absence of customs within the EU results in a second feature, namely the fact that importers do not pay VAT when the goods clear customs but are responsible for self-reporting their imports when their VAT return is due, i.e. with a time lag between the moments of purchase and reporting. The combination of these two features resulted in the emergence of Missing Trader Intra-Community (MTIC) fraud (de la Feria, 2019, p. 173). The most basic incarnation of MTIC fraud involves a tax-registered firm importing goods from another EU country (VAT-free), subsequently selling it domestically (charging VAT) and disappearing before remitting the VAT collected from its client. MTIC schemes are essentially an extreme version of sales under-reporting in which no sales are reported. More sophisticated versions of the MTIC scheme exist, involving buffer (legal) companies and repeating the fraud above in a loop. Sometimes, the missing trader does not simply disappear but files for bankruptcy before remitting VAT. For a comprehensive description of each scheme — which all have in common that a firm goes missing before remitting the VAT is legally owed — see European Commission (2018, pp. 9-14).

Besides fraud involving a missing trader, the Reverse Charge Mechanism on cross-border transactions and the absence of customs give rise to two other fraud strategies. First, firms may disguise domestic sales as intra-community supplies, allowing them to charge VAT without remitting it to the tax authorities (since within-EU exports are VAT-free). Second, firms may make false claims for refunds by importing goods from another EU Member State — on which VAT is due domestically — and claiming, however, that the purchase was domestic — in which case they would have paid VAT to the seller and would be entitled to a refund. Doing so allows them to pocket a refund for a purchase on which they actually did not pay VAT. However, MTIC fraud is believed to be quantitatively more important (Lamensch and Ceci, 2018).

---

[9] It applies to international transactions in almost all VAT systems around the world.

3.3.3 The Reverse Charge Mechanism on domestic transactions as a solution to fraud

The European Commission has been aware of the possible fragility of the system since its creation, and has been battling fraud ever since. One particular reform aiming at stifling MTIC-type fraud is the extension of the reverse charge mechanism to selected *domestic* transactions: by removing the ability of the missing trader to charge VAT on its domestic sales, the fraudulent scheme is effectively neutralized (in theory at least) as fraudsters cannot collect VAT anymore. The disguise of domestic transactions as cross-border activity (and conversely) to unduly gain VAT monies also become ineffective.

The application of the DRCM to selected domestic supplies became progressively available to Member States over time on a optional basis and subject to notification to the European Commission. Importantly, Member States must subsequently file a report on the apparent effectiveness of the measure. In other words, implementing the DRCM is costly and endogenous to the level of fraud. The scope of products eligible for DRCM widened over time, starting from construction work, supply of waste and immovable property in 2006 (art. 199 of the VAT Directive) to trade in emission permits in 2010 (art. 199a of the VAT Directive) and to a wider range of goods including mobile telephones, game consoles, laptops, supplies of cereals and raw and semi-finished metals in 2013 (modification of art. 199a of the VAT Directive). The DRCM has also been granted to some countries for selected goods and services on the basis of art. 395 of the VAT Directive (de la Feria, 2019).[10] In some instances, the DRCM is only imposed to transactions whose value exceed a threshold so as to reduce the compliance and administrative costs to small businesses.

Whilst it effectively removes the ability of the missing trader to collect VAT (and run away with it), the DRCM presents its own challenges that are currently heatedly debated among lawmakers in the EU. The main problem is the removal of the chain of incremental cross-checks and tax collection that make VAT so attractive (see section 3.2), transforming the VAT system into a Retail Sales Tax in which the tax is only col-

---

[10] Note that many of the supplies specified in art. 199 and 199a are services, and therefore beyond the scope of the present study that only focuses on tradable goods. There are two other and more recent legal bases for applying the DRCM. Since, 2013 art. 199b VAT Directive allows the quick application of the DRCM "to combat sudden and massive fraud liable to lead to considerable and irreparable financial losses". Since, 2019 art. 199c VAT Directive allows the temporary application of a Generalized DRCM to all domestic transactions.

lected at the last stage of the chain.[11] Furthermore, MTIC fraud may be displaced to other products that are not subject to the DRCM. The next section details the theoretical link between MTIC fraud and the trade reporting gaps.

3.4 VAT fraud and the reporting Gap

To measure the importance of pre-reform fraud, I propose to investigate how bilateral trade asymmetries — i.e. the difference between reports of a given trade flow by the exporting and the importing countries, that I denote *reporting gaps* — react to the implementation of the DRCM. Intuitively, fraud schemes involving a missing trader or disguising domestic transactions as cross-border trade will result in trade asymmetries. In a nutshell, the idea of the empirical exercise is to observe how trade asymmetries change once the ability to fraud is removed, and thereby estimate the importance of fraud prevailing prior to the reform. In this section, I formally define the measure of trade asymmetries and establish how they are expected to change in response to the DRCM.

3.4.1 The reporting gaps

The reporting gaps are constructed using mirror statistics, i.e. information on the same gross flow that is reported twice: once by the importer (as an import) and once by the exporter (as an export). The gaps are computed in terms of value and quantities. Formally:

$$X\ \text{gap}_{iepym}^{log} = \log(\text{exports } X_{ipym}^{e}) - \log(\text{imports } X_{epym}^{i}), \tag{3.1}$$

where $X \in \{\text{value}, \text{quantity}\}$, i denotes the importing country, e denotes the exporting country, p the product and y and m refer to the year and month of observation. The superscripts denote who reports the trade flow and subscripts designate the trade partner. The reporting gaps as calculated in (3.1) can be read as the size of the gap expressed in terms of the value of reported imports. In Bussy (2020), I describe at length what factors may influence the reporting gaps and this discussion is largely relegated there, except for specifics relevant to the current study.

---

[11] Naritomi (2019) shows that evasion appears to be particularly important at the last stage of the chain.

3.4.2 Theoretical relationship between VAT fraud and the reporting gaps

Any type of fraud scheme whereby one side has an incentive to report a transaction whereas the other does not will result in a disparity in trade reports and therefore in the reporting gaps. I consider in turn the three types of fraud introduced in section 3.3.2 focusing on their effect on the reporting gaps and how the DRCM affects this relationship. I explicitly derive testable predictions that will be empirically investigated.

3.4.2.1 Missing trader fraud

An illustration of the most basic incarnation of this fraud can be found in figure 3.B.1. The importer purchases supplies from abroad and sells them on domestically, charging VAT, and disappears before remitting VAT to the authorities. As a result, the transaction is reported on the exporter's side, but not on the importer's side. This results in an increase in the reporting gap as defined in (3.1). Once the DRCM is implemented, it is not possible for the importer to charge VAT to the local firm anymore. As MTIC-related transactions stop, the reporting gap are predicted to decrease — reported imports remain unchanged since the missing trader never reported, and reported exports decrease.

Prediction 1. *Upon implementation of the DRCM in the importing country, the reporting gaps associated to those products subject to it decrease.*

Another prediction immediately follows: MTIC fraud may be displaced to other products which have similar characteristics. Typically, MTIC fraud is concentrated among high-value and low-weight/volume products - such as precious metals, computer chips and laptops.

Prediction 2. *As MTIC fraud is displaced to other products in response to the implementation of the DRCM in the importing country, the reporting gaps associated with those products increase.*

3.4.2.2 Imports disguised as domestic purchases

In this type of fraud, firms import goods from another EU Member State, on which VAT is due domestically and based on their own report. They claim, however, that the purchase was domestic, in which case they would have paid VAT to the seller and would be entitled to a refund. Doing so allows them to pocket a refund for a purchase on which they actually did not pay VAT at the time of purchase. An illustration is provided in figure 3.B.2. This behaviour impacts the reporting gaps positively, as the exporter reports truthfully whereas the importer does not report its imports. With the DRCM in place, VAT on domestic purchases cannot be refunded, because these are effectively treated as cross-border transactions: they report it, pay VAT on it and get the refund simultaneously.

Prediction 3. *When the DRCM is implemented in the importing country, firms become unable to unduly claim refunds by disguising imports as domestic purchases. They report their imports truthfully and the reporting gaps decrease.*

3.4.2.3 Domestic sales disguised as exports

Firms claim to make sales of goods to another EU member, which are VAT free. In reality they sell the goods within their country, charging VAT and pocketing the VAT amount instead of remitting it to the tax authorities. Figure 3.B.3 contains an illustration of this strategy. Since exports are artificially inflated whilst reported imports are not impacted, the reporting gaps rise. With the DRCM in place, the firm pretending to export cannot charge VAT domestically on the actual sale, and therefore cannot fraud in this fashion.

Prediction 4. *When the DRCM is implemented in the exporting country, firms cannot collect VAT on domestic sales, eliminating the incentive to disguise them as exports and resulting in a decrease in the reporting gaps.*

3.4.3 Underlying assumptions

To summarize, the reporting gaps are expected to react unambiguously to the implementation of the DRCM both in the importing and exporting countries. Regarding MTIC fraud, the underlying assumptions are that the missing trader does not report its purchases, whereas the exporter does. In all variants of the MTIC scheme, the exporter is either an honest firm or an accomplice that must look like as honest firm. It is therefore reasonable to assume that it will dutifully report the transaction. Since the missing trader disappears before filing its VAT return, it is very likely that the transaction is not reported on that side of the reporting gaps.

However, not all firms are obliged to report their transactions for statistical purposes. As explained in details in Bussy (2020, section 3.2.2), only firms whose volumes of imports and exports exceed specified thresholds report to a system called Intrastat, on the basis of which official trade statistics are produced. These thresholds are set such that 97% of total exports and 93% of total imports are accounted for, and given the skewed distribution of firms' size, many smaller firms are exempt from reporting. In 2017, an average of 12.6% and 20.2% of VAT-registered traders were providers of statistical information for imports and exports, respectively (Eurostat, 2017). If the exporter is exempt from reporting, no discrepancy will appear in the reporting gaps as neither side is reporting. In that sense, this empirical strategy may miss fraud perpetrated by smaller firms that do not show up in trade statistics.

3.5 Data

In this section, I describe the data on the DRCM events and on the reporting gaps. The data on all DRCM events was hand-collected, mostly from national laws and European Commission documents. Details can be found in appendix 3.C. Trade data underlying the reporting gaps is from Eurostat.

3.5.1 Sample definition

I identified 54 instances — defined as a country and time — where the DRCM was applied to selected products, denoted $v = 1, \ldots, 54$. A list of the events can be found in table 3.A.1: they span 24 EU countries and over the years 2004-2019. There are two samples at a monthly frequency: the *importer sample* in which the treatment (i.e. the implementation of the DRCM) takes place in the importing country, and the *exporter sample* in which it takes place in the exporting country. They cover trade between the country in which the reform takes place and its partners over a time window spanning 12 months before and after each event. A graphical representation can be found in figure 3.B.4. The event time is denoted $t \in [-A, B]$, where $A = 12, B = 11$ and the event takes place at $t = 0$ by convention. An observation is a product p at the 8-digit level of the Combined Nomenclature (CN), traded between importing country i and exporting country e in event v at a given event time t. Note that each observation is observed at a given calendar date (month and year ym).[12] Overall, the samples contain between 36 and 37 million observations and a summary of the relative (aggregated)

---

[12] There is a one-to-one mapping between event number-event time (vt) and calendar year-month (ym). Therefore, an observation can be equivalently defined as $eipym$ or $eipvt$.

TABLE 3.1: OBSERVATIONS PER GROUP

	Exporter sample				Importer sample			
	With overlap		Without overlap		With overlap		Without overlap	
Group	N	# *p*	N	# *p*	N	# *p*	N	# *p*
Control	36,755,945	13,676	29,653,654	13,657	36,380,729	13,927	31,133,306	13,916
Treatment	266,177	971	176,654	856	274,433	1,056	189,376	971

Note: Number of observations and unique products (#*p*) for the control and treatment groups, summed across all events. The treatment group is composed of observations to which the DRCM applies. This table only contains observations where the reporting gap in terms of value is non-missing. The columns without overlap refer to observations for which DRCM reforms do not overlap with each other over time within a country.

sizes of the control and treatment groups can be found in table 3.1. Although there is no perfect overlap of events within countries (in which case the events would count as one), 13 out of the 54 events have a partial overlap with each other — i.e. the DRCM was implemented several times *within the same country* less than 2 years apart. Last, very few observations for which the DRCM was introduced and then abolished within 12 months of the introduction have been dropped from the sample.

3.5.2 Reporting gaps

The reporting gaps are defined at the CN 8-digit level — the finest grain available from Eurostat — since the treatment is often defined at such a dis-aggregated level.[13] Descriptive statistics of the reporting gaps can be found in table 3.2. The mean gap hovers between 2% and 5% across samples, and can be extremely large. The positive mean indicates that on average, reported exports exceed reported imports. This may be due to the fact that Intrastat thresholds for reporting imports are generally higher than in the case of exports (cf. section 3.5.3). Figure 3.B.5 displays the histograms of the reporting gaps, as well as the mean and median gaps over months. As previously discussed in Bussy (2020), the gaps are strongly cyclical. For robustness, the regressions are also estimated dropping the top and bottom percentiles of the reporting gaps.

3.5.3 Other data

Data on Intrastat thresholds was hand-collected and is freely available here (last accessed 19/05/2020). Data on the standard VAT rate, which applies to the vast majority

---

[13] For instance, when the DRCM applies to waste and scraps of metal, it may only target 8-digit products. As an illustration, waste and scrap of manganese is product code 8111 00 11 and the 6-digit code 8111 00 does not only cover waste and scrap, but other manganese products to which the DRCM does not apply.

TABLE 3.2: DESCRIPTIVE STATISTICS OF THE REPORTING GAPS

	Importer sample							
	N	Mean	Std	Min	25%	50%	75%	Max
Value gap, in log	36,655,162	0.04	1.93	-17.20	-0.58	0.02	0.76	17.20
Quantity gap, in log	32,652,504	0.05	2.10	-18.85	-0.69	0.01	0.87	19.14

	Exporter sample							
	N	Mean	Std	Min	25%	50%	75%	Max
Value gap, in log	37,022,122	0.03	1.94	-16.41	-0.60	0.01	0.75	17.20
Quantity gap, in log	33,043,192	0.02	2.11	-18.15	-0.69	0.00	0.83	18.11

Note: The importer sample is that in which the events occur in the importing country, and the exporter sample that in which the events occur in the exporting country.

of products, come from the European Commission. Data on the VAT gaps come from Barbone et al. (2013) (for years 2000-2011), Poniatowski et al. (2018) (for 2012) and Poniatowski et al. (2019) (for 2013-2017). The *Burden of customs procedure* and *Efficiency of customs* indices describing customs quality are from the World Bank, and the *Corruption* index is provided by the International Country Risk Guide (ICRG).

3.6 Empirical strategy

In this section, I detail the regression equations that will be estimated, as well as discuss the underlying identification assumptions and threats thereto.

3.6.1 Regression equations

The general strategy is a difference-in-difference approach where the treatment is the implementation of the DRCM. For each instances of the DRCM, treated products are those to which the DRCM applies and the control group is composed of all the other products to which the DRCM does not apply.

Static setup The static specification of the difference-in-difference regression is

$$\text{Gap}_{jt} = \gamma_j + \delta_t + \beta D_{jt} + \mathbf{FE} + \epsilon_{jt}, \tag{3.2}$$

where the index j denotes a *unit*: a product p traded between importing country i and exporting country e in event v, i.e. $j \equiv (i, e, p, v)$. The dependent variable Gap_{jt} is a measure of the reporting gap. $t \in [-A, B]$ denotes the event time, which spans A periods before the treatment and B periods after the treatment. In most specifications, $A = 12, B = 11$ such that the event window spans 12 months on either side of the event time. γ_j and δ_t are unit and event time fixed effects, respectively. D_{jt} is a dummy that takes value 1 if observation j is subject to the DRCM at time t.[14] **FE** is a vector of fixed effects, which either include calendar time (*ym*), calendar time-product (*ymp*), calendar time-importer and calendar time-exporter (*ymi* and *yme*), calendar time-country pair (*ymie*) fixed effects. Event and calendar times differ because events are scattered over time and across countries and products. Given the seasonality in the reporting gaps, including calendar time fixed effects is desirable. In light of predictions 1, 3 and 4, $\hat{\beta}$ — the estimate of β — is expected to be negative when (3.2) is estimated off the importer or exporter samples.

Dynamic specification In order to check that the parallel pre-trend assumption holds as well as study the dynamic effects of the treatment, a dynamic version of (3.2) is estimated as follows:

$$\text{Gap}_{jt} = \gamma_j + \delta_t + \sum_{k=-A}^{B} \beta_k D_{jt}^k + \mathbf{FE} + \epsilon_{jt}, \tag{3.3}$$

where $D_{jt}^k = \mathbb{1}[t = e_j + k]$ is a dummy indicating that the event took place k periods ago — e_j is the date at which unit j has been treated. The regression admits A leads (pre-trends) and B lags (dynamic effects) of the treatment event dummy. $\beta_k = 0, \forall k < 0$ is the necessary condition to rule out differential pre-trends between treatment and control groups. In light of predictions 1, 3 and 4, the estimated dynamic effects $\hat{\beta}_k, \forall k > 0$ are expected to be negative when (3.3) is estimated off the importer or exporter samples.

3.6.2 Identification

First, I describe the typical timing of a DRCM reform and its implications for identification. Second, I discuss the two main threats to identification in the empirical strategy above: reverse causality and omitted variable bias.

Timing Art. 199 and 199a VAT Directive allow Member States to apply the DRCM to specified good on a voluntary basis. Any implementation must be notified to the European Commission. Since the DRCM reforms are subject to publication in the national laws and therefore subject to political scrutiny that takes time, the reform

---

[14] The estimated β in (3.2) is the same as would be obtained by interacting dummies for the treatment group and the post period, conditional on including j and t fixed effects.

is not a surprise to fraudsters. I define the announcement date as the moment at which the DRCM is passed into law — firms may know of the reform even before that, as project of laws are made public before their final approval and publication. On average, the DRCM reform was announced 70 days prior to its implementation across all events. As a consequence, there is a possible anticipation bias if fraud ceases before the actual reform date. Hence the importance of estimating equation (3.3), in which one might detect this issue if one observes $\hat{\beta}_k \neq 0$ for k immediately before the event time.[15]

Reverse causality Reverse causality concerns arise if fraud causes the implementation of the DRCM, which we know is precisely what happens in reality. There are two dimensions of potential reverse causality in the present case. First, a cross-sectional dimension: those products to which the application of the DRCM is legally permitted are fraud-prone (art. 199 and 199a VAT Directive). Second, a time dimension: the significant administrative costs associated with the application of the DRCM must be justified by strong suspicions of fraud for it to be worth the while of Member States. It is therefore possible that only when fraud becomes important enough over time do countries implement the DRCM. The high frequency of the data allows to look within a restricted time frame within which one can be confident that the reporting gaps react to the DRCM and not the other way around, which should mitigate the time dimension the reverse causality. However, since the DRCM applies to goods most vulnerable to fraud, the estimated coefficients in regressions (3.2) and (3.3) can only be interpreted as the effect of the DRCM in this fraud-prone environment. It is for instance impossible to infer the size of cross-border VAT fraud in general from these estimates.

Omitted variable bias As explained in Bussy (2020, section 3.2), the reporting gaps encompass many factors unrelated to evasion or fraud, such as transportation costs — that only appear in reported imports — and asymmetries in the Intrastat declaration thresholds between the importing and the exporting countries — that can result in one side declaring, whereas the other does not. Generally, the reporting gaps are functions

---

[15] The announcement date always falls within the event window. The earliest a DRCM reform was announced was 254 days prior to implementation.

of VAT fraud and other potentially confounding factors:

$$\text{Gap}_{jt} = f(\text{VAT fraud}_{jt}, \chi_{jt}) \tag{3.4}$$

where χ_{jt} denotes the components of the gap that do not pertain to VAT fraud. The challenge of using the reporting gaps to identify VAT fraud becomes apparent in this expression. Not controlling for the determinants of χ_{jt} in the regression would result in an omitted variable bias, unless χ_{jt} is uncorrelated with D_{jt}. Identification thus relies on fixed effects and controls to account for the components of χ_{jt} that may be correlated with the timing or scope of the reform.

Since treatment is at the level of a product-country pair-event (*iepv*), many fixed effects can be included in the regressions in an attempt to control for potentially confounding elements of χ_{jt}. For instance, changes in the Intrastat reporting threshold or in the corporate income tax rate in the exporting country, which are common to all products, would be absorbed by an exporter-time (*et*) fixed effect. Besides, limiting the study to a narrow time window around the event decreases the risk of unobserved factors confounding the results.

Other factors There are two other factors that are likely to result in underestimating MTIC fraud as measured by the reaction of the reporting gaps to the DRCM. First, as mentioned in section 3.4.3, traders with lower annual trade volumes are not required to report for statistical purposes. Although the levels of thresholds are controlled for in regressions (via country-time fixed effects), the reporting gaps would not react to the implementation of the DRCM if all exporters were exempted from reporting the transaction associated with the fraud. While there is anecdotal evidence that some MTIC fraud takes place in the form of smaller transactions to avoid detection, statistical agencies have noticed discrepancies in reported trade. For instance, the HMRC in the United Kingdom actually *corrects* aggregated trade statistics for MTIC activities (HMRC, 2019).[16] Furthermore, more evolved versions to the MTIC scheme involve

---

[16] HMRC (2019, p. 9) writes:

> Because the recording of EU trade data is based on Intrastat declarations from VAT registered businesses, the way in which the fraudulent transactions are reported means that any exports relating to the fraud are reported, but imports (arrivals) relating to the fraud are not. As a consequence, UK import (arrival) statistics would be under reported without an adjustment for MTIC.

The adjustment is made in division 99 of the CN classification, as part of the monthly non-response estimate. The method used to estimate MTIC in the UK is confidential.

a network of firms to which misreporting (at other stages than the missing trader) would attract unwanted attention. In conclusion, in is likely that part of MTIC fraud will be undetected in the reporting gaps.

Second, the application of the DRCM is sometimes limited to transactions whose value exceed a certain threshold in order to minimize the compliance costs to small businesses. If fraud took place in the form of very low volume transactions only the DRCM would be ineffective. However, this threshold is typically low ($\leq$ EUR 10,000), uncommon — applied in 7 member states for selected goods as of 2014 (European Commission, 2014) — and seems unlikely to be binding.

3.7 Results

In this section, I detail the results of the study, starting with a visual inspection of the treatment effect in the raw data, followed by estimates of the difference-in-difference exercise laid out in section 3.6. Last, I provide estimates of the level of fraud removed by the DRCM.

3.7.1 Visual inspection

Predictions 1, 3 and 4 all imply a negative impact of the implementation of the DRCM on the reporting gaps, irrespective of whether the reform takes place in the importing country — which prevents MTIC schemes and profits from disguising imports as domestic purchases — or in the exporting country — in which case misreporting domestic sales as exports is not profitable anymore. Figures 3.B.6 to 3.B.10 and 3.B.11 to 3.B.15 plot the mean reporting gap (in terms of value) over time across products in the treatment group (units j that are subject to the DRCM) and the control group (the rest of the units) in the importer and exporter samples. Under the aforementioned predictions, one would expect the mean gap in the treatment group to decrease relative to that in the control group from the reform time onward.

These figures offer two main insights. First, it is not the case that the reporting gaps systematically decrease after the DRCM is implemented. Although it clearly does in some events (e.g. events 5, 31, 32, 42 or 45 when the DRCM is implemented in the importing country), it is unclear in most instances and some events display the opposite pattern (i.e. the gaps in the treatment group increase relative to the control group, e.g. in events 17, 41 or 51 when the DRCM is implemented in the importing country). Second, the mean gap is considerably more volatile in the treatment group relative to the control group. This is because there are much fewer observations in the treatment group.

Tables 3.A.2 to 3.A.5 provide the number of observations and products in the treatment groups, as well as the mean gap pre- and post-reform for each event separately (when the reform takes place in the importing and exporting country). These tables confirm that the control groups generally dwarf the treatment groups. The difference-in-difference in reporting gap means (in value terms) is also provided — calculated using the raw data.[17] In 32 (31) out of 54 events in which the reform takes place in the importing (exporting) country, the DiD estimate is negative. Although this represents the majority of the cases, there is a significant fraction of events in which the reporting gaps associated with the treated products do not decrease relative to the control group.

3.7.2 Baseline results

Static estimates Estimates of equation (3.2) in the importer and the exporter samples (i.e. when the DRCM takes place in the importing and the exporting country, respectively) can be found in table 3.3. Columns (1) to (5) contain the estimates from regressions with increasingly stringent sets of fixed effects.

Upon implementation of the DRCM in the importing country, the reporting gaps experience a significant decrease of 2.5% to 3.5% depending on the fixed effects included in the regressions.[18] The coefficients are relatively stable across specifications, suggesting that the fixed effects do not pick up significant confounding omitted variables. The number of observations in column (5) is lower as including *ymkp* fixed effects restricts the sample to events that have a time overlap across countries, i.e. a product (p) traded with partner country (k) at a given time (ym) must be observed across more than one reform country (c). The results are consistent with predictions 1 and 3. On the one hand, the DRCM prevents MTIC fraud such that reported exports that were associated with the fraud disappear whilst reported imports remain constant (as the missing trader never reported the imports), implying a decrease in the reporting gaps. On the other hand, firms that disguised imports as domestic purchases

---

[17] The difference-in-difference (DiD) is calculated as DiD $= (\overline{\text{Gap}}_{cont.,pre} - \overline{\text{Gap}}_{cont.,post}) - (\overline{\text{Gap}}_{treat.,pre} - \overline{\text{Gap}}_{treat.,post})$.

[18] The estimates of β are semi-elasticities in this case. The implementation of the DRCM is associated with a $100\beta\%$ change in the reporting gaps.

TABLE 3.3: BASELINE STATIC DIFFERENCE-IN-DIFFERENCE: VALUE GAP

	Importer sample				
	Dependant variable: Value gap, log				
	(1)	(2)	(3)	(4)	(5)
DRCM (D_{jt})	-0.0349***	-0.0248**	-0.0314***	-0.0302***	-0.0280**
	(-3.51)	(-2.07)	(-3.15)	(-3.03)	(-2.22)
Adjusted R^2	0.478	0.483	0.479	0.479	0.508
Within R^2	0.000	0.000	0.000	0.000	0.000
Observations	36,115,872	35,957,790	36,115,854	36,115,584	29,952,923
	Exporter sample				
	Dependant variable: Value gap, log				
	(1)	(2)	(3)	(4)	(5)
DRCM (D_{jt})	-0.0101	0.00272	-0.0102	-0.00990	-0.0102
	(-1.14)	(0.21)	(-1.14)	(-1.10)	(-0.66)
Adjusted R^2	0.484	0.491	0.485	0.485	0.502
Within R^2	0.000	0.000	0.000	0.000	0.000
Observations	36,477,528	36,314,706	36,477,528	36,477,290	27,287,522
t, j FE	✓	✓	✓	✓	✓
ymp FE		✓			
ymc FE			✓		✓
ymk FE			✓		
$ymck$ FE				✓	
$ymkp$ FE					✓

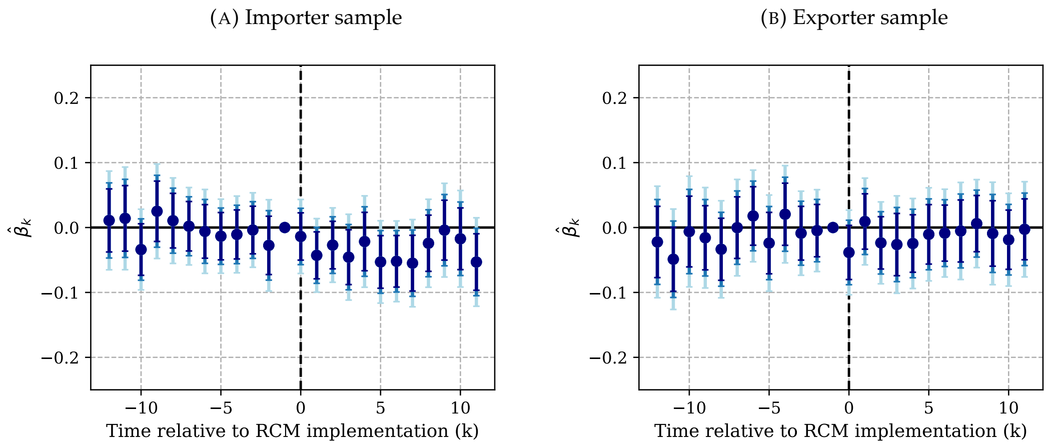
Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Fixed effects (FE): event time (t), unit ($j \equiv iepv$) calendar time (ym), product (p), country in which the reform takes place ($c = i$ in the importer sample, and $c = e$ in the exporting sample), partner country ($k = e$ in the importer sample, and $k = i$ in the exporting sample). The dependent variable is the reporting gap in terms of value, in logs as specified in expression (3.1).

to unduly pocket the VAT rebate cannot profit from this scheme anymore. Reported imports increase as they report truthfully, lowering the reporting gaps. I attempt to translate these point-wise elasticities into levels in section 3.7.4.

The implementation of the DRCM in the exporting country does not impact the reporting gaps significantly. Point estimates suggest a decrease of the reporting gaps of about 1%, but this effect is not precisely estimated. The negative impact on the gaps is consistent with prediction 4: once the DRCM is applied, firms cannot benefit from disguising domestic sales as exports, a strategy that previously allowed them to collect VAT without remitting it to the authorities. Reported exports decrease and so do the reporting gaps.

Overall, the results are consistent with the predictions. The fact that the effect of the DRCM when implemented in the exporting country is smaller in magnitude and not significant is also telling and consistent: not only are *two* types of frauds neutralized when the DRCM is implemented in the importing country (MTIC fraud and disguised imports), but it also is expected that MTIC fraud is the quantitatively most important

FIGURE 3.1: BASELINE DYNAMIC DIFFERENCE-IN-DIFFERENCE: VALUE GAP



Note: Sequences of $\hat{\beta}_k$ from equation (3.3), estimated on the importer and exporter samples with the following fixed effects: $j, t, ymc, ymkp$. The coefficient in the period immediately preceding the implementation of the DRCM has been normalized to 0. 99%-, 95%- and 90%-confidence intervals are depicted in different colour shades.

fraud Lamensch and Ceci (2018).

Dynamic estimates One concern with the estimates above is the potential existence of differential pre-trends. For instance, if the reporting gaps associated with treated products were on a downward trend over the whole event window (including prior to the implementation of the DRCM), that would be picked up as a negative $\hat{\beta}$ in regression (3.2). However, this would be caused by differential pre-trends and not by the DRCM intervention. Estimates of the series of β_k in equation (3.3), with the strictest set of fixed effects $(y, t, ymc, ymkp)$, can be found in figure 3.1. First, the parallel pre-trend assumption does not appear to be violated in either sample, since none of the $\hat{\beta}_k$, $\forall k < 0$ are significant. Second, while the RMC does not have any significant impact when implemented in the exporting country, it appears to have a negative impact in the months following its implementation in the importing country: all point estimates of the dynamic effects are negative. These findings echo the results from the static specification in table 3.3. However, only few dynamic effects post-DRCM are significant, and it seems that they revert back towards zero after the 9th month post-treatment, in contrast to the predicted permanent decrease in the reporting gaps. These patterns are robust to the inclusion of all the combination of fixed effects listed in table 3.3.

Summary and robustness checks Overall, the findings are in line with the predictions from section 3.4, although they lack statistical significance in the dynamic specification. The reporting gaps react negatively to the implementation of the DRCM in the importing country, consistent with the elimination of MTIC fraud and imports disguised as domestic purchases. When the DRCM is applied in the exporting country, the effect on the reporting gaps are negative although smaller and not significant. This is also consistent with the supposition that fraudulent disguise of domestic sales as exports is quantitatively less important than MTIC fraud.

The results are robust to balancing the sample (see table 3.A.6), to removing event overlaps (see table 3.A.7), to dropping extreme values in terms of the reporting gaps (see table 3.A.8) and to controlling for the levels of VAT main rates and Intrastat thresholds in both countries (when the fixed effects allow, see table 3.A.9).[19] Throughout these additional specifications, the results are stable and the magnitudes of the effects of the DRCM on the reporting gaps tend to be slightly larger than the baseline in table 3.3. Last, as shown in table 3.A.10, the results are robust to controlling for the implementation of the DRCM in the partner country — in table 3.3, this is only the case in column (5) with the inclusion of $ymkp$ fixed effects.[20]

3.7.3 Further findings

In this section, I provide a series of additional results based on specifications akin to (3.2) and (3.3), but exploiting additional heterogeneity across countries or products.

3.7.3.1 Reported traded quantities also react to the DRCM

The predictions in section 3.4 should hold for the reporting gaps expressed in terms of values as well as quantities, since for each fraud type the implementation of the DRCM modifies the firms' incentives to report whole transactions — unlike in the case of tariff evasion where the traders may have an incentive to lie on the unit value as opposed to the quantity, e.g. if the goods are differentiated and it is harder for customs to determine the true price (Javorcik and Narciso, 2008). For instance in the case of MTIC fraud, the DRCM removes the ability of the importer to carry out the fraud and all the trade generated by the scheme prior to the reform is predicted to disappear post-

---

[19] A tiny fraction of products are not subject to the main rate, but to reduced rates. It is not easy to know which products are subject to reduced rates as it varies from country to country and over time and it is not systematically documented in terms of CN codes. However, since the vast majority of products are subject to the main rate, ignoring this imprecision should not change the results.

[20] Controlling for reforms in the partner country is especially important since the set of products to which the DRCM can be applied is identical across Member States, which increases the risk of a reform affecting a trade flow simultaneously in the origin and destination countries.

reform, affecting the trade value and traded quantities alike.[21] The estimates from regressions (3.2) and (3.3) using the reporting gaps in terms of quantities as dependent variable can be found in table 3.A.11 and figure 3.B.16, respectively. The results are almost identical to the baseline, suggesting that firms change their reports of whole transactions when the DRCM is implemented.

3.7.3.2 The effects of the DRCM are heterogeneous across events

Up until now, the estimation was based on all cases pooled together, yielding an average treatment effect across all events. However, it is probable that the DRCM reform did not have the same effect in each episode. I estimate equation (3.2) separately for each event and plot the resulting estimates in figure 3.2.[22] There is indeed significant heterogeneity across events. First, whereas it is expected that $\hat{\beta}_v < 0$, $\forall v$, there are some events in which the DRCM is associated with an increase in the reporting gaps, contrary to all predictions. Second, the effect of the DRCM on the reporting gaps is not significant in many cases when estimated on an event-by-event basis. Third, there is considerable variation in terms of the magnitudes of the estimated β, which range from -0.63 to 0.53. In comparison, the point estimates in the baseline specification were around -0.03 and -0.01 when the reform takes place in the importing and exporting countries, respectively.

The interpretation of positive values of $\hat{\beta}_v$ in light of the predictions from section 3.4 is challenging. Those are events in which the reporting gaps of the treated products increased relative to those of other goods. A naive approach is to consider that no fraud was eliminated by the introduction of the the DRCM in these instances since the reporting gaps did not decrease. A more prudent path is to consider the existence of other factors that impacted the reporting gaps at the time of reform, resulting in the observed increase. Given the high number of coefficients estimated, positive estimates could be the result of innocuous measurement errors.[23] These omitted factors might

---

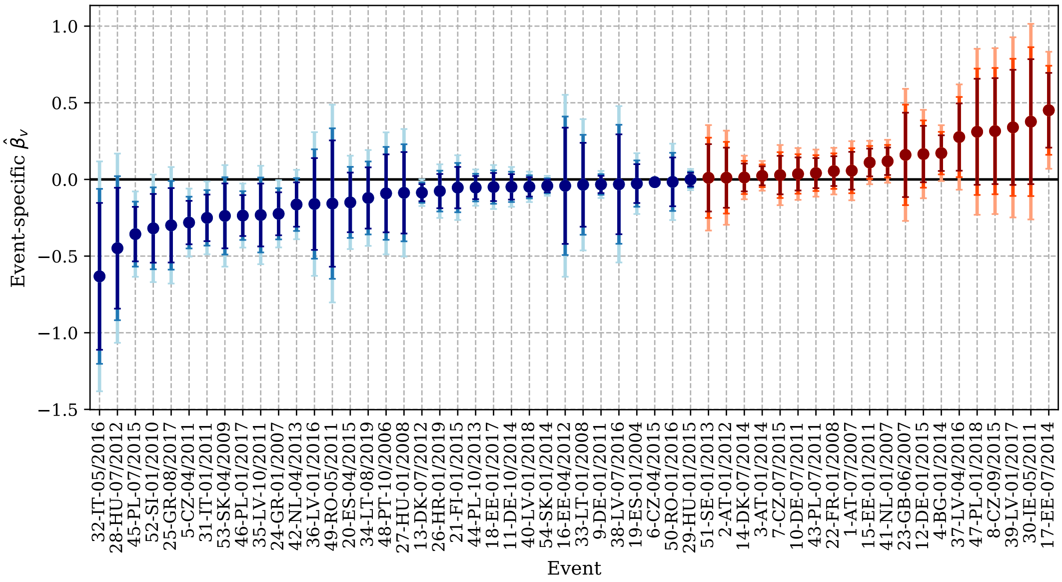
[21] A similar logic applies to the disguise of imports as domestic purchases and of domestic sales as exports: firms misreport the whole transaction prior to the reform and stop post-reform, affecting traded values and quantities alike.

[22] Alternatively, I estimate equation (3.2) on the whole sample interacting the treatment dummy (D_{jt}) with a dummy for the country in which the reform takes place. The results are very similar. Note that when estimating (3.2) event by event, it is not possible to include *ymkp* fixed effects since at a given point in time (*ym*), a good (*p*) traded with a partner (*k*) appears only once in the event-specific sample.

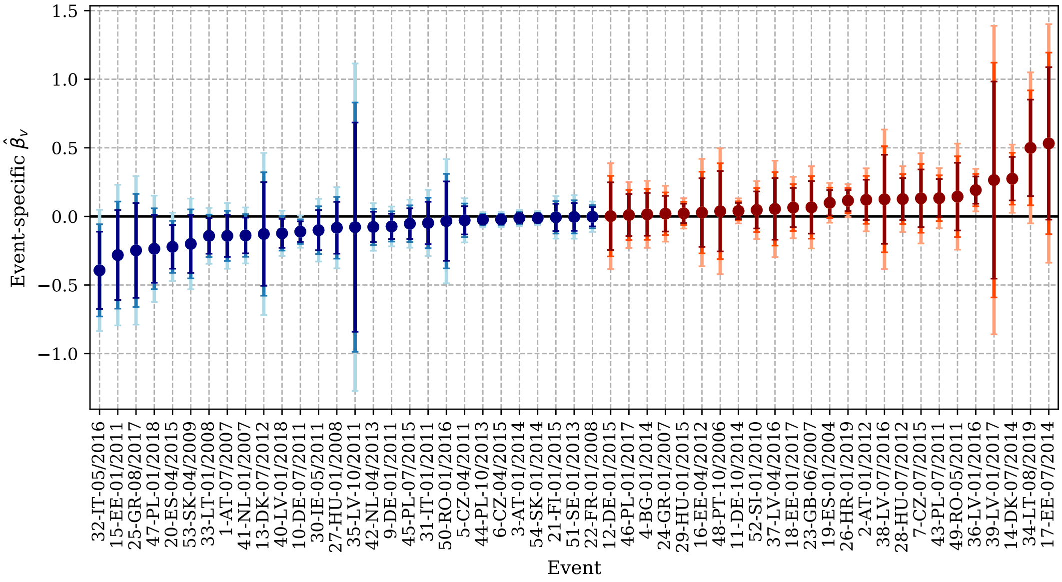
[23] However, this cannot result from classical measurement errors (i.e. errors uncorrelated with the latent true variables). In that case, measurement errors in the dependent variable (which is more likely in the present case) leave OLS unbiased and consistent, whereas measurement errors in the independent variables result in an attenuation bias but cannot cause a switch in sign.

FIGURE 3.2: EVENT BY EVENT DIFFERENCE-IN-DIFFERENCE: VALUE GAP

(A) Importer sample



(B) Exporter sample

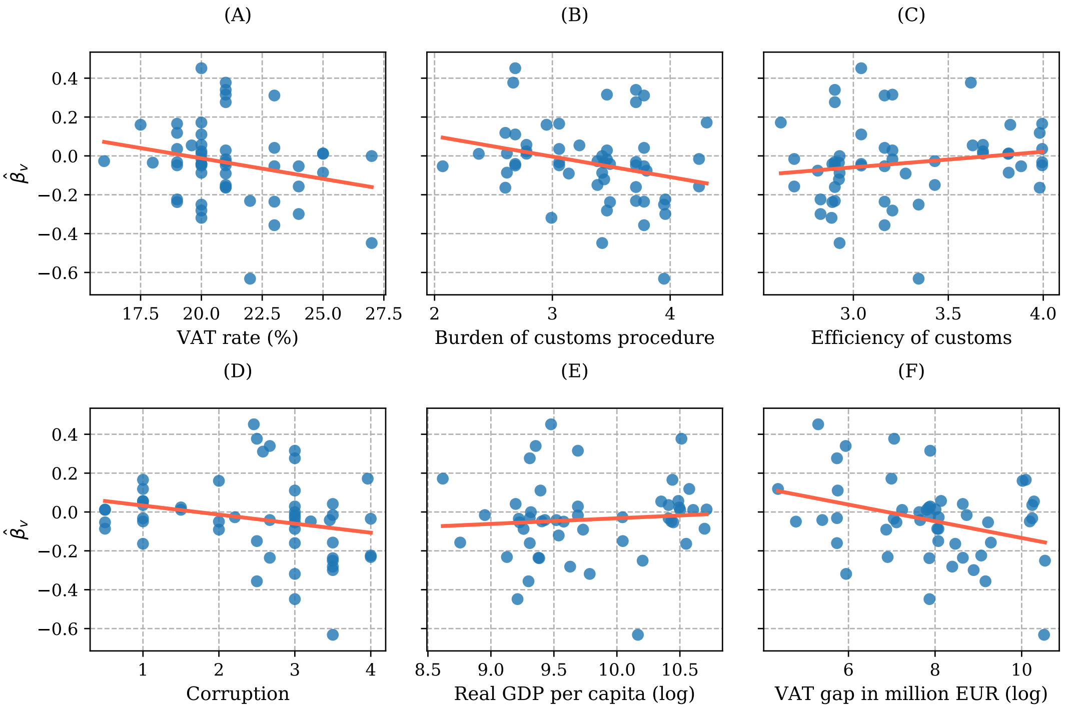


Note: Event-specific $\hat{\beta}$ from equation (3.2), obtained by estimating the equation separately for each event with fixed effects $j, t, ymck$. 99%-, 95%- and 90%-confidence intervals are depicted in different colour shades: blue (red) when point estimates are negative (positive).

however bias results, not only in episodes with a positive estimated coefficient but also in events where the effect of the reform has the predicted sign.

From here on, I focus on the effects of the DRCM on reported imports by the reform country - i.e. on the importer sample. Overall, it appears that a few events might be influencing the results significantly. There are about 12 events in which the estimated effect of the DRCM on the reporting gaps is negative, of high magnitude and significant. On the other end of the spectrum, there are 9 events in which $\hat{\beta}_v$ is large

Figure 3.3: Correlations between event-specific coefficients and other variables



Note: A dot represents an event. The VAT rate is the standard rate. *Burden of customs procedure* is an index ranging from 1 (extremely efficient) to 7 (extremely inefficient). *Efficiency of customs* is and index ranging from 1 (low) to 5 (high). *Corruption* is an index ranging from 0 (low corruption) to 6 (high corruption).

and positive — yet significant in fewer cases. There are no clear patterns between the event-specific $\hat{\beta}_v$ and obvious observables: neither the reform country and dates, the type of treated products or the number of observations are systematically linked to the sign or magnitude of $\hat{\beta}_v$. The reform was announced well in advance (up to 254 days) in some of the events in which $\hat{\beta}_v > 0$. The average number of days between announcement and implementation is 61 across events where $\hat{\beta}_v < 0$ and 86 across events where the DRCM is associated with an increase in the reporting gaps. If fraud ceases before the reform takes place, the effect on the reporting gaps is expected to be less pronounced. Yet this does not justify the positive signs of the estimated coefficients. Furthermore, a reform does not have symmetrical effects on the reporting gaps associated with flows where the country is importing or exporting.[24]

The event-specific coefficients are mildly correlated with other variables of interest in intuitive ways. As a reminder, the more negative $\hat{\beta}_v$ is, the stronger the reaction of the reporting gap and the higher the inferred level of fraud pre-reform. As depicted in figure 3.3, the estimated coefficients are more negative for higher levels of VAT rates

---

[24] In other words, there is no correlation between the $\hat{\beta}_v$ estimated off the importer and exporter samples, apart from the two extremes: EE-07/2014 and IT-05/2016, where the reform is associated with similar effects when the country is an importer or an exporter.

(more details follow in section 3.7.3.4), when customs are of lower quality, for higher levels of corruption and finally when the VAT gap — a measure of non-compliance in VAT collection — is higher. There is no apparent link between the level of real GDP per capita in the importing country and the estimated coefficient.[25] Customs quality and corruption indices are meant to capture overall quality of institutions, and VAT fraud — as proxied by $\hat{\beta}_v$ — appears larger in countries that score poorly in these dimensions. Similarly, VAT fraud seems more important in countries with a high VAT gap — further analysis follows in section 3.7.4. Overall, although the estimated coefficients correlate intuitively with other observables, positive values of $\hat{\beta}_v$ cannot be easily justified.

3.7.3.3 Stronger effects when the control group is limited to similar goods

As evident from tables 3.A.2 to 3.A.5, the control groups generally dwarf the treatment groups as only a small subset of products at the 8-digit level are subject to the DRCM. One possible concern is that given the volatility in the reporting gaps, having the entire set of products in the control group introduces additional noise that could mask the true effect of the DRCM reforms. The set of fixed effects included in regressions largely absorbs this variation in the data, but even in the specifications with the strictest set of fixed effects the adjusted R^2 hoovers around 0.5, suggesting that significant variation is left. In order to lower the size of the control group, I restrict it to products from the same section in the Combined Nomenclature (CN) as the treated units. There are 21 sections grouping products according to the broad industry to which they belong.[26] The underlying motivation for this *ex-ante* restriction is that similar products might be subject to similar shocks over time, which fixed effects may be better able to absorb and neutralize in the restricted sample.

The results are displayed in table 3.A.12 and figure 3.B.17, for the static and dynamic specifications, respectively. The estimates are similar to those in the baseline, but of higher magnitude when the DRCM is implemented in the importing country. Whereas the reporting gaps were found to decrease by between 2.5% and 4.5% upon implementation of the DRCM in the baseline case, they decrease by 4.5% to 7.7% when the control group is restricted to products in the same section as the treated units. Re-

---

[25] The correlations depicted in figure 3.3 remain similar in sign and magnitude conditioning on the effect of the DRCM being negative, i.e. dropping all events in which $\hat{\beta}_v < 0$.

[26] For instance, section I is "Live animals; animal products", section II is "Vegetable products", section XI is "Textiles and textile articles", etc.

TABLE 3.4: STATIC DIFFERENCE-IN-DIFFERENCE: VALUE GAP — WITH SECTION FIXED EFFECTS

	Importer sample		
	Dependant variable: Value gap, log		
	(1)	(2)	(3)
DRCM (D_{jt})	-0.0534***	-0.0497***	-0.0626***
	(-3.73)	(-2.85)	(-2.96)
Adjusted R^2	0.474	0.478	0.500
Within R^2	0.000	0.000	0.000
Observations	30,876,640	30,710,753	24,709,260
	Exporter sample		
	Dependant variable: Value gap, log		
	(1)	(2)	(3)
DRCM (D_{jt})	-0.00598	0.00947	0.00366
	(-0.48)	(0.51)	(0.16)
Adjusted R^2	0.483	0.490	0.494
Within R^2	0.000	0.000	0.000
Observations	29,358,056	29,182,204	20,524,362
t, j FE	✓	✓	✓
ymp FE		✓	
ymkp FE			✓
ymcks FE	✓	✓	✓

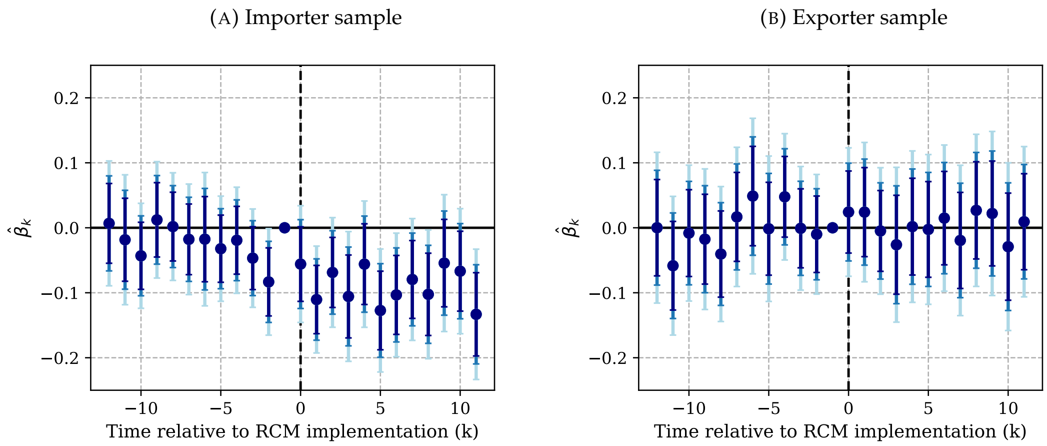
Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Fixed effects (FE): event time (t), unit ($j \equiv iepv$) calendar time (ym), product (p), country in which the reform takes place ($c = i$ in the importer sample, and $c = e$ in the exporting sample), partner country ($k = e$ in the importer sample, and $k = i$ in the exporting sample), CN product section (s). The dependent variable is the reporting gap in terms of value, in logs as specified in expression (3.1).

stricting the sample in this fashion also has drawbacks, as (i) information is lost, (ii) the definition of the control group is arbitrary. Alternatively, I add section (s)-calendar time (ym)-country pair (ck) fixed effects without restricting the control group to goods of the same section and find very similar results: a decrease in reporting gaps by 5% to 6.3% with more pronounced dynamic effects as reported in table 3.4 and figure 3.4. There are signs that non-parallel pre-trends may be present in panel (A) of figure 3.4 as the point estimates slowly decrease over time and even become significant in one pre-reform period. This problem is only present when *ymkp* fixed effects are included, which limits the number of observations in the sample.

3.7.3.4 The effects of the DRCM does not strongly depend on the VAT rate

The marginal benefit of misreporting one Euro of a transaction for fraud purposes is directly proportional to the VAT rate. In each fraud scheme detailed in section 3.3.2, fraudsters illegally gain VAT on the misreported transaction. It is therefore possible that fraud might be more prevalent in high-VAT countries. To test this hypothesis, I estimate equation (3.2) interacting the treatment dummy with the rate of VAT. Given

FIGURE 3.4: DYNAMIC DIFFERENCE-IN-DIFFERENCE: VALUE GAP — WITH SECTION FIXED EFFECTS



Note: Sequences of $\hat{\beta}_k$ from equation (3.3), estimated on the importer and exporter samples with the following fixed effects: $j, t, ymkp, ymcks$. The coefficient in the period immediately preceding the implementation of the DRCM has been normalized to 0. 99%-, 95%- and 90%-confidence intervals are depicted in different colour shades.

the above reasoning, the coefficient on the interaction should have a negative sign as the effect of the DRCM is expected to be exacerbated by high VAT rates. Fixed effects should absorb other country-specific characteristics correlated with the levels of VAT rates.

The results are in table 3.A.13. Although the coefficient on the interaction has the expected sign in most specifications in the importer sample, it is only significant in the first two columns — with less rich fixed effects. The interaction does not have the expected sign and is never significant in the exporter sample. Overall, the evidence does not point to a strong impact of the rates of VAT on fraud as measured by the reaction of the reporting gaps to the reform.

3.7.3.5 The DRCM does not appear to shift fraud to other products

From the reports following the implementation of the DRCM by Member States, it appears that fraud shifted to other products not subject to the DRCM in some cases.[27] Although there is mixed anecdotal evidence for this shift, I check whether I find signs of it in the data. Equation (3.2) is estimated including a dummy that takes value 1 post-treatment for goods to which fraudsters are likely to turn when the DRCM is implemented. I consider three classes of such products. First, goods with high unit value, which are known to be attractive for fraud purposes as they maximize value for a given traded quantity. Second, other goods to which the DRCM is applicable

---

[27] See sections 3.2.3 and 3.2.4 of European Commission report *COM/2018/0118 final*.

according to the VAT Directive, but to which it was *not* applied in that specific reform episode. These goods are in the law precisely because they have been identified as fraud-prone. Third, goods within the same HS heading (HS 4-digit group) as some Member States reported a shift of fraud to very similar goods falling just outside of the scope of the DRCM.

The results are displayed in table 3.A.14, with a focus the importer sample. There is no evidence of movements in the reporting gaps of goods with high unit value. However, it appears that the reporting gaps associated with goods to which the DRCM can be applied (but is not) *decrease*. Speculatively, this could be consistent with fraudsters moving their activities to other goods or another country, lowering fraud levels even for goods to which the DRCM did not apply. Last, the reporting gaps associated with products within the same HS heading weakly increase (only significantly in some specifications), consistent with fraud shifting to these goods at the time of reform. These mixed results are not conclusive. Furthermore, fraud may have shifted to other goods not considered here or to other countries.

3.7.4 Levels of fraud

So far, the focus was on the reaction of the reporting gaps to the implementation of the DRCM, interpreted as evidence of fraudulent activity ceasing. The point estimates give an idea of the importance of fraud : reported exports decrease by around 3% relative to reported imports when the DRCM is implemented in the importing country, consistent with the prediction that the DRCM removes the incentive to engage in strategies resulting in unreported imports. This section provides tentative estimates of the *monetary amounts* of fraud eliminated by the DRCM, or in other words the revenue losses incurred by Member States prior to the implementation of the DRCM. The analysis focuses on the effects of the reform on reported imports (importer sample). All monetary amounts are in 2015 real EUR.

3.7.4.1 Revenue losses prevented by the DRCM

The basic idea is to estimate the amount of VAT that was collected and not remitted by fraudsters pre-reform based on the point estimates indicating how the reporting gaps change when the DRCM is implemented. For each episode of the reform (v), I

TABLE 3.5: DESCRIPTIVE STATISTICS OF FRAUD AMOUNTS ACROSS EVENTS

	N	Mean	Std	Min	25%	50%	75%	Max
Trade (m. EUR)	52	1,918.18	3,052.77	6.70	115.99	554.56	1,712.94	15,093.22
VAT rate (%)	52	21.08	2.25	16.00	20.00	21.00	22.25	27.00
Common $\hat{\beta}$	52	-0.03	0.00	-0.03	-0.03	-0.03	-0.03	-0.03
Event-specific $\hat{\beta}_v$	32	-0.15	0.15	-0.63	-0.24	-0.09	-0.04	-0.00
Fraud (m. EUR) — $\hat{\beta}_v$	32	50.41	84.54	0.06	2.68	9.21	59.37	412.40
Fraud (% rev.) — $\hat{\beta}_v$	32	0.21	0.30	0.00	0.02	0.17	0.26	1.37
Fraud (% VAT gap) — $\hat{\beta}_v$	31	0.91	1.16	0.01	0.09	0.45	1.29	4.32
Fraud (m. EUR) — $\hat{\beta}$	52	11.72	18.09	0.04	0.74	4.10	11.07	86.03
Fraud (% rev.) — $\hat{\beta}$	52	0.06	0.12	0.00	0.01	0.02	0.05	0.66
Fraud (% VAT gap) — $\hat{\beta}$	50	0.52	1.66	0.00	0.04	0.11	0.31	11.46

Note: v refers to the event number. Data: importer sample. Amounts of trade and fraud, when not expressed as fractions, are in 2015 million EUR. (% rev.) means expressed in percentage of VAT revenues in the year of reform. (% VAT gap) means expressed in percentage of the VAT gap in the year of reform. All amounts are annual values.

estimate the amount of fraud pre-reform as follows:

$$\text{Fraud amount}_v^{pre} = \underbrace{\text{trade volume subject to the DRCM}_v \times \mid \hat{\beta}_v \mid}_{\text{tax base associated with fraud}} \times \text{VAT rate}_v, \quad (3.5)$$

where the VAT rate is the standard one.[28] The trade volume subject to the DRCM is measured as 12 times the average monthly reported exports of the treated goods over the 12 months prior to the reform — i.e. total trade over the year prior to the reform when no observations are missing.[29] Trade flows are measured by reported exports because both in the case of MTIC and imports disguised as domestic purchases, the fraud amounts are not reported by the importer and therefore do not appear in reported imports.[30] The last element of (3.5) is the estimated effect of the reform, $\hat{\beta}_v$. Two options are considered: (i) the value estimated on all events simultaneously from table 3.3, so $\hat{\beta}_v = -0.03, \forall v$; (ii) event-specific estimates from panel (A) in figure 3.2, *setting them to zero when positive*. Neither solution is completely satisfactory: (i) ignores the important heterogeneity across events and instead gives an average treatment effect; (ii) implicitly assumes that an observed increase in the reporting gaps at the time of reform ($\hat{\beta}_v > 0$) is tantamount to absence of fraud. I report the results using both options.

---

[28] The list of products to which reduced rates can be applied can be found in Appendix I VAT Directive. The only possible overlap between these goods and those to which the DRCM is applicable is agricultural inputs. After further inspection, it appears that in only two instances (events 7 and 8 in CZ) the goods to which the DRCM applied might have been subject to a reduced rate (of 15% instead of 21%). Switching one for the other barely changes the results.

[29] I use 12 times the monthly average in case some observations are missing, i.e. I implicitly impute the average trade value for missing observations.

[30] I also recompute calculations using reported imports, or the average of reported imports and exports. The results do change and the resulting estimated evasion is somewhat smaller.

The resulting estimated amount of fraud removed by the implementation of the DRCM for each event separately can be found in table 3.A.15, and summary statistics across events are in table 3.5. All amounts are expressed annually and levels are in million 2015 EUR. Several results spring to the eyes: First, whether event-specific or a common estimated effect of the DRCM is used in (3.5) makes a large difference. This originates from the fact that $\hat{\beta}_v$ estimated event-by-event are generally larger in magnitude (conditional on being negative) than the common $\hat{\beta}$. Furthermore, the trade volume subject to the DRCM was important in several events in which $\hat{\beta}_v$ is particularly negative. To give an idea of how the results differ depending on this choice, summing fraud in the year leading to the reform across all events amounts to 1.6 billion EUR using event-specific coefficients, whereas it sums up to a mere 610 million EUR based on a common coefficient value. However, a striking conclusion arises no matter which option is chosen: the estimated fraud levels represent only a tiny fraction of VAT revenues, between 0.06% and 0.2% on average across events. Furthermore, it only represents $0.5\% - 0.9\%$ of the VAT gap, which is meant to capture VAT non-compliance.[31] In section 3.7.4.3, I describe more thoroughly the relationship between the VAT gaps and the DRCM. The instance of the reform that removed the most fraud (using event-specific estimates $\hat{\beta}_v$) is in Poland in 2015 (event $v = 45$) when the DRCM was applied to a wide range of products including mobile phones, laptops and supplies of raw and semi-finished metals.

3.7.4.2 Difficulties in extrapolating aggregate fraud levels

The calculations above need to be interpreted with appropriate care. First, the fraud levels are those in the year prior to the reform in these countries and for the treated goods only. In other words, it is not implied that fraud levels were the same in every year before the reform, or that implementing the DRCM to other goods will have a similar effect. Actually, given the selection bias whereby the DRCM is only applicable to goods known to be subject to fraud, it is very likely that applying the DRCM elsewhere would remove much less fraud. Second, fraud in a country may not decrease by the full amount removed by the reform if fraudsters shift their activities to other goods within the same country — although I have found no evidence of this.

---

[31] I also computed the fraud levels using a common $\hat{\beta} = 0.055$, averaging coefficients across specifications in table 3.4 (which includes CN section fixed effects). Total fraud in the year leading to reform is ca. EUR 1.17 billion. Removed fraud represents 0.1% of VAT revenues and 0.94% of VAT gaps on average across events.

To illustrate the selection bias and the dangers of extrapolating fraud levels associated with transactions that are not subject to the reform, I consider two extreme hypothetical scenarios: (i) fraud is limited to the products subject to the reform, and (ii) fraud levels are uniform across all products. The truth lies somewhere in-between. Under assumption (i), all fraud is removed by the reform and the estimates based on equation (3.5) capture the total level of fraud (associated with the schemes considered in this study). Under assumption (ii), I estimate total fraud by replacing the first term in (3.5) by the total trade volume between the reform country and its partners in the year prior to the reform - as opposed to the trade volume subject to the DRCM only.[32] Across all events, total fraud in the year leading to the reform under (ii) would represent 145% of VAT revenues and 676% of the VAT gaps on average. These unrealistic figures confirm that the magnitudes of fraud directly removed by the DRCM cannot be extrapolated to all other cross-border transactions.

3.7.4.3 VAT Gap and the DRCM

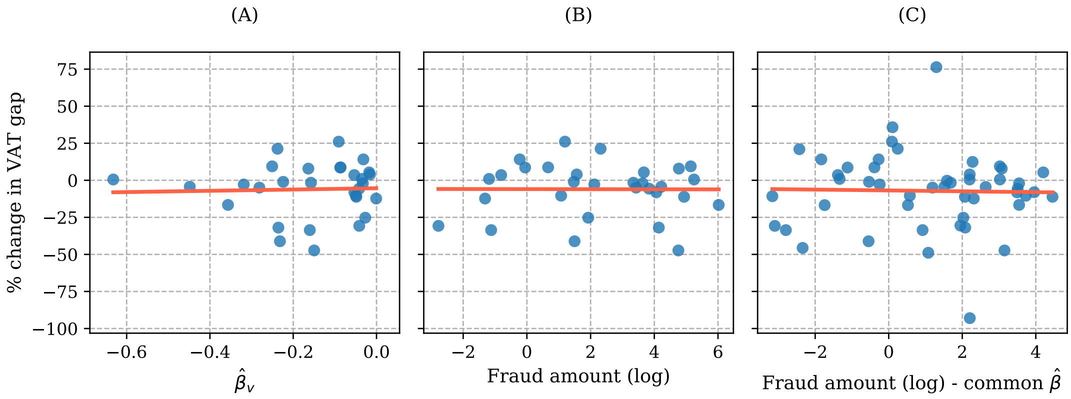
One natural question is whether the VAT gaps computed by tax authorities of the Member States react to the implementation of the DRCM. Fraud being a component of the VAT gap, a decrease following the reform is expected. The analysis is limited by the restricted number of observations available as the VAT gaps are only available at the yearly frequency. Furthermore, as eluded to in section 3.3, the VAT gaps can be caused by many factors in addition to fraud, such as non-compliance, legitimate errors by firms in filling their VAT returns, bankruptcies and insolvencies (Poniatowski et al., 2019).

Graphical inspection of the time series of the VAT gaps at the times of reform, displayed in figures 3.B.18 and 3.B.19, does not suggest a strong incidence of the DRCM. Out of the 51 events for at which time the VAT gap is available, there are 30 cases in which the VAT gap decreases in the year of implementation (relative to the previous year), but in 19 of out of these cases the VAT gap already decreased in the previous year. Furthermore, there is no systematic relation between either the magnitude of $\hat{\beta}_v$ or the value of fraud as calculated in (3.5) and changes in the VAT gap (see figure 3.5).

---

[32] On average across all events, the trade volume subject to the DRCM represents 0.17% of total trade.

FIGURE 3.5: CHANGE IN VAT GAPS VERSUS ESTIMATED FRAUD



Note: A dot represents an event. Scatter plot of the percentage change in the VAT gap between the year of implementation and the year before (y-axis) versus: (A) event-specific effect of the DRCM; (B) log of the fraud amounts as calculated in (3.5) in million real EUR, using event-specific $\hat{\beta}_v$; (C) log of fraud amounts as calculated in (3.5) in million real EUR, using the common $\hat{\beta}$ from table 3.3.

3.7.5 Discussion

Policy focus Although the VAT gaps encompass many factors and are prone to measurement errors, the absence of movement in the VAT gaps at the time of reforms together with the lack of correlation between the VAT gaps and the estimated magnitudes of fraud are puzzling, especially given how important MTIC fraud is believed to be by tax authorities in the EU. Even if my estimates underestimate fraud (cf. explanations in section 3.6.2), the implementation of the DRCM should be noticeable in the VAT gaps if MTIC is responsible for large tax revenue losses. Since Member States that implemented the DRCM generally found the measure to be effective, either the scale of MTIC fraud is smaller than previously thought or the VAT gaps suffer from measurement errors precluding quantitative conclusions.[33]

Cost-benefit analysis of the DRCM From the outset, I would like to emphasize that precise calculations and interpolations are impossible. The discussion below merely evokes some elements to fuel the policy debate. The implementation of the DRCM to selected products comes at a cost both for the regulator and for firms. An attempt at estimating the costs incurred by firms to comply with the DRCM was made by European Commission (2014). They estimate that the reverse charge mechanism results in additional compliance costs amounting to 0.13% of turnover based on a survey of

---

[33] The views of Member States on the DRCM and its effectiveness are summarized in European Commission Report *COM(2018) 118 final*.

34 businesses across six Member States.[34] They estimate that 2.17% of Gross Value Added was generated from activities subject to the DRCM in 2014, concluding that the compliance cost amounted to EUR 323 million in that year (European Commission, 2014, p. 62). On the other hand, my estimates suggest that fraud worth between EUR 423 million and EUR 739 million (depending on the choice of $\hat{\beta}$) was eliminated by the DRCM annually as of 2014.[35] From this simple comparison, the additional tax revenues outweigh the compliance costs. There are however other costs to the implementation of the DRCM borne by authorities that are not taken into account here, and perhaps other benefits than the boost in tax revenues — e.g. the removal of firms operating in the shadow economy selling at unreasonably low prices. However, this exercise suggests that costs and benefits might not differ by orders of magnitude.

3.8 Conclusion

The introduction of the Reverse Charge Mechanism to domestic transactions eliminates a class of VAT fraud schemes involving cross-border transactions whereby fraudsters collect VAT without remitting it to the tax authorities. When engaging in fraud, firms misreport their transactions, which in turn affects the country's trade statistics. This paper proposes an estimation of the amount of fraud removed by the reforms based on the movement of discrepancies between reports by the importing and exporting countries of the same trade flows around the reform time. I find that in the months following the reform, the reporting gaps associated with treated goods *imported* by the reform country decrease by around 3% on average relative to goods not subject to the DRCM. This is consistent with the removal of MTIC fraud and of the fraudulent disguise of imports as domestic purchases. The effects of the reform vary greatly across episodes. Overall, I estimate that these reforms removed up to EUR 1.6 billion in fraud annually, a large figure in absolute terms, but tiny compared to VAT revenues.

3.9 Bibliography

Baldwin, R, "EU VAT Fraud," *Vox: Research-based Policy Analysis and Commentary from Leading Economists*, 2007.

---

[34] It is unclear how precise this estimate is given the very low number of sampled firms.

[35] This is assuming that the fraud pre-reform would have continued at the same intensity over the years in the absence of the DRCM. Note that I compare the compliance costs of firms subject to the DRCM to fraud removed by the DRCM. Inferring fraud levels for other goods is thus not necessary here and therefore not an issue.

Barbone, Luca, Misha Belkindas, Leon Bettendorf, Richard M Bird, Mikhail Bonch-Osmolovskiy, and Michael Smart, "Study to quantify and analyse the VAT Gap in the EU-27 Member States," *CASE Network Reports*, 2013, (116).

Benzarti, Youssef and Alisa Tazhitdinova, "Do Value-Added Taxes Affect International Trade Flows? Evidence from 30 Years of Tax Reforms," Technical Report, National Bureau of Economic Research 2019.

Bussy, Adrien, "Corporate Tax Evasion: Evidence from International Trade," Working Paper 2020.

CASE, "Assessment whether the introduction of VAT reverse charge mechanism on 1 October 2013 for certain steel products limited VAT fraud with respect to the trade in these goods and whether the business development environment in the industry under changes has improved (trad.)," 2015.

de la Feria, Rita, "The New VAT General Reverse-Charge Mechanism," *EC Tax Review*, 2019, *28* (4), 172–175.

Demir, Banu and Beata Javorcik, "Trade policy changes, tax evasion and Benford's law," *Journal of Development Economics*, 2020, p. 102456.

European Commission, *Assessment of the application and impact of the optional 'Reverse Charge Mechanism' within the EU VAT system* 2014.

_ , "The concept of Tax Gaps. Report III: MTIC Fraud Gap estimation methodologies," November 2018.

Eurostat, "National requirements for the Intrastat system. 2018 Edition," 2017.

Farhad, Mohammad, Michael Jetter, Abu Siddique, and Andrew Williams, "Misreported Trade," 2018.

Ferrantino, Michael J, Xuepeng Liu, and Zhi Wang, "Evasion behaviors of exporters and importers: Evidence from the US–China trade data discrepancy," *Journal of International Economics*, 2012, *86* (1), 141–157.

Fisman, Raymond and Shang-Jin Wei, "Tax rates and tax evasion: evidence from "missing imports" in China," *Journal of political Economy*, 2004, *112* (2), 471–496.

Gemmell, Norman and John Hasseldine, "Taxpayers' behavioural responses and measures of tax compliance 'gaps': A critique and a new measure," *Fiscal Studies*, 2014, *35* (3), 275–296.

Gradeva, Katerina, "VAT Fraud in Intra-EU Trade," *Goethe University, Frankfurt*, 2014.

Grásgruber, Miloš, Milena Otavová, Pavel Semerád et al., "Impacts of the application of the reverse charge mechanism of the value added tax," *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, 2013, *61* (7), 2133–2141.

HMRC, "Overseas Trade Statistics Methodology Paper," *HM Revenue & Customs –Trade Statistics*, October 2019.

Hopland, Arnt O and Robert Ullmann, "Pushing the wrong buttons: VAT evasion by misclassification of meal consumption type," *European Accounting Review*, 2019, pp. 1–23.

Javorcik, Beata S and Gaia Narciso, "Differentiated products and evasion of import tariffs," *Journal of International Economics*, 2008, *76* (2), 208–222.

_ and _ , "WTO accession and tariff evasion," *Journal of Development Economics*, 2017, *125*, 59–69.

Keen, Michael and Joel Slemrod, "Optimal tax administration," *Journal of Public Economics*, 2017, *152*, 133–142.

_ and Stephen Smith, "VAT fraud and evasion: What do we know and what can be done?," *National Tax Journal*, 2006, pp. 861–887.

Kellenberg, Derek and Arik Levinson, "Misreporting trade: Tariff evasion, corruption, and auditing standards," *Review of International Economics*, 2019, *27* (1), 106–129.

Lamensch, Marie and Emanuele Ceci, "VAT fraud: Economic impact, challenges and policy issues," Technical Report 2018.

Levin, Jörgen and Lars M Widell, "Tax evasion in Kenya and Tanzania: Evidence from missing imports," *Economic Modelling*, 2014, *39*, 151–162.

Mishra, Prachi, Arvind Subramanian, and Petia Topalova, "Tariffs, enforcement, and customs evasion: Evidence from India," *Journal of public Economics*, 2008, *92* (10), 1907–1925.

Naritomi, Joana, "Consumers as tax auditors," *American Economic Review*, 2019, *109* (9), 3031–72.

Pomeranz, Dina, "No taxation without information: Deterrence and self-enforcement in the value added tax," *American Economic Review*, 2015, *105* (8), 2539–69.

Poniatowski, Grzegorz, Mikhail Bonch-Osmolovskiy, José Ma Duran-Cabré, Alejandro Esteller-Moré, and Adam Śmietanka, "Study and Reports on the VAT Gap in the EU-28 Member States: 2018 Final Report," *Available at SSRN 3272816*, 2018.

_ , _ , _ , _ , and _ , "Study and Reports on the VAT Gap in the EU-28 Member States: 2019 Final Report," 2019.

Stiller, Wojciech and MA Heinemann, "Reverse-Charge-Mechanism-An effective measure against VAT fraud in Germany and Austria?," *BETRIEBSWIRTSCHAFTLICHE FORSCHUNG UND PRAXIS*, 2019, *71* (2), 177–+.

Stoyanov, Andrey, "Tariff evasion and rules of origin violations under the Canada-US Free Trade Agreement," *Canadian Journal of Economics/Revue canadienne d'économique*, 2012, *45* (3), 879–902.

Zídková, Eliška Čejková-Hana, "The Impact of Specific Reverse Charge on Waste and Scrap on Tax Revenues in the Czech Republic," 2019.

Appendix 3.A Additional tables

TABLE 3.A.1: LIST OF INSTANCES OF APPLICATION OF THE DRCM

N	Country	Implementation date	Goods
1	AT	01/07/2007	199.1(d)
2	AT	01/01/2012	199a(c), 199a(d)
3	AT	01/01/2014	199a(h), 199a(j)
4	BG	01/01/2014	199a(i)
5	CZ	01/04/2011	199.1(d)
6	CZ	01/04/2015	199a(c), 199a(d), 199a(h), 199a(j)
7	CZ	01/07/2015	199a(i)
8	CZ	01/09/2015	199a(i)
9	DE	01/01/2011	199.1(d), 199a(j)
10	DE	01/07/2011	199a(c), 199a(d)
11	DE	01/10/2014	199a(h), 199a(j)
12	DE	01/01/2015	199a(j)
13	DK	01/07/2012	199.1(d)
14	DK	01/07/2014	199a(c), 199a(d), 199a(h)
15	EE	01/01/2011	199.1(d)
16	EE	01/04/2012	199a(j)
17	EE	01/07/2014	199a(j)
18	EE	01/01/2017	199a(j)
19	ES	01/01/2004	199.1(d)
20	ES	01/04/2015	199a(j), 199a(c), 199a(h)
21	FI	01/01/2015	199.1(d)
22	FR	01/01/2008	199.1(d)
23	GB	01/06/2007	199a(c), 199a(d)
24	GR	01/01/2007	199.1(d)
25	GR	01/08/2017	199a(c), 199a(h)
26	HR	01/01/2019	199a(j)
27	HU	01/01/2008	199.1(d)
28	HU	01/07/2012	199a(i)
29	HU	01/01/2015	199a(j)
30	IE	01/05/2011	199.1(d)
31	IT	01/01/2011	199a(c), 199a(d)
32	IT	02/05/2016	199a(h)
33	LT	01/01/2008	199.1(d), EXCEPTION
34	LT	01/08/2019	199a(c), 199a(h), EXCEPTION
35	LV	01/10/2011	199.1(d)
36	LV	01/01/2016	EXCEPTION
37	LV	01/04/2016	199a(c), 199a(d), 199a(h)
38	LV	01/07/2016	199a(i)
39	LV	01/01/2017	199a(j)
40	LV	01/01/2018	199a(h), 199a(j), EXCEPTION
41	NL	01/01/2007	199.1(d)
42	NL	01/04/2013	199a(c), 199a(d), 199a(h)
43	PL	01/07/2011	199.1(d)
44	PL	01/10/2013	199.1(d)
45	PL	01/07/2015	199a(j), 199a(h), 199a(c)
46	PL	01/01/2017	199a(j), 199a(d)
47	PL	01/01/2018	EXCEPTION
48	PT	01/10/2006	199.1(d)
49	RO	31/05/2011	199a(i)
50	RO	01/01/2016	199a(c), 199a(d), 199a(h)
51	SE	01/01/2013	199.1(d)
52	SI	01/01/2010	199.1(d)
53	SK	01/04/2009	199.1(d), 199a(j)
54	SK	01/01/2014	199a(c), 199a(d), 199a(i), 199a(j)

Note: The *Goods* columns lists the articles on which basis the DRCM was implemented. 199.1(d): waste and scraps; 199a(c): mobile phones; 199a(d): integrated circuit devices; 199a(h): laptops and game consoles; 199a(i): cereals and industrial crops; 199a(j): supplies of raw and semi-finished metals. EXCEPTION indicates that the DRCM applies to goods not specified in articles 199 or 199a VAT Directive.

TABLE 3.A.2: DRCM INSTANCES IN IMPORTING COUNTRIES (CONTINUES IN NEXT TABLE)

Case	Grp	N	# prod	$\overline{\text{Gap}}_{pre}$	$\overline{\text{Gap}}_{post}$	Δ Means	DiD	Mean Freq.
1: AT, 7.2007	0	734438	9329	-0.021	-0.03	0.0092		12
	1	3013	50	-0.23	-0.12	-0.11	0.12	12
2: AT, 1.2012	0	813492	8933	-0.048	-0.066	0.018		13
	1	3076	17	-0.2	-0.11	-0.091	0.11	16
3: AT, 1.2014	0	836405	7940	-0.051	-0.052	0.0007		14
	1	34043	475	-0.059	-0.05	-0.0094	0.01	12
4: BG, 1.2014	0	466654	7079	-0.17	-0.12	-0.053		10
	1	484	31	-0.49	-0.18	-0.31	0.26	4.8
5: CZ, 4.2011	0	873891	8790	-0.13	-0.14	0.0091		13
	1	2622	61	0.22	-0.0081	0.22	-0.21	10
6: CZ, 4.2015	0	912203	7701	-0.16	-0.14	-0.012		14
	1	67024	597	-0.18	-0.19	0.016	-0.028	14
7: CZ, 7.2015	0	989709	8305	-0.16	-0.14	-0.011		14
	1	1225	18	-0.18	-0.34	0.16	-0.17	10
8: CZ, 9.2015	0	997909	8325	-0.16	-0.14	-0.012		14
	1	17	1	-0.3	-0.024	-0.27	0.26	8.5
9: DE, 1.2011	0	1367204	8749	0.13	0.079	0.047		15
	1	10604	64	-0.014	-0.098	0.084	-0.036	14
10: DE, 7.2011	0	1386606	9522	0.1	0.089	0.012		15
	1	5425	17	-0.28	-0.25	-0.03	0.043	18
11: DE, 10.2014	0	1524439	8635	0.1	0.094	0.0097		16
	1	8992	82	0.2	0.1	0.094	-0.084	13
12: DE, 1.2015	0	1543770	8634	0.096	0.095	0.0014		16
	1	2309	30	0.16	0.41	-0.26	0.26	11
13: DK, 7.2012	0	759730	8311	0.12	0.083	0.035		13
	1	430	21	0.27	0.54	-0.27	0.3	9.3
14: DK, 7.2014	0	801465	7779	0.08	0.054	0.026		14
	1	3274	19	-0.029	-0.062	0.033	-0.0071	15
15: EE, 1.2011	0	479818	6869	0.076	0.089	-0.013		11
	1	163	16	0.36	0.62	-0.27	0.25	5.3
16: EE, 4.2012	0	513534	7236	0.09	0.093	-0.0035		11
	1	23	4	0.52	-0.28	0.8	-0.81	2.9
17: EE, 7.2014	0	557447	6985	0.092	0.13	-0.037		12
	1	511	12	0.41	0.92	-0.51	0.47	9.8
18: EE, 1.2017	0	612671	7130	0.15	0.16	-0.0021		12
	1	12152	237	0.18	0.12	0.058	-0.061	10
19: ES, 1.2004	0	775646	9440	0.066	0.11	-0.04		14
	1	4443	70	0.16	0.084	0.08	-0.12	12
20: ES, 4.2015	0	1024537	8667	0.067	0.092	-0.025		14
	1	1708	11	0.3	0.27	0.036	-0.061	15
21: FI, 1.2015	0	662480	7456	0.084	0.099	-0.016		13
	1	133	9	0.0047	-0.054	0.059	-0.074	8.3
22: FR, 1.2008	0	1155608	9170	-0.039	-0.029	-0.011		15
	1	2685	41	0.027	-0.0033	0.03	-0.041	12
23: GB, 6.2007	0	933170	9674	0.075	0.11	-0.037		13
	1	769	3	-0.051	0.057	-0.11	0.072	13
24: GR, 1.2007	0	500130	8575	0.037	0.04	-0.0026		10
	1	366	21	-0.04	-0.012	-0.028	0.025	5.5
25: GR, 8.2017	0	626788	8123	0.1	0.11	-0.0043		11
	1	1386	4	0.34	-0.033	0.37	-0.37	16
26: HR, 1.2019	0	736217	7667	-0.15	-0.13	-0.014		13
	1	1782	20	-0.19	-0.22	0.029	-0.043	14
27: HU, 1.2008	0	630027	7126	0.086	0.15	-0.066		12
	1	557	15	-0.071	0.064	-0.14	0.069	8.7

Note: Grp (0/1) refers to treatment and control groups, i.e. those trade flows subject to the DRCM (1) or not (0). N is the number of observations in each group, and # prod refers to the number of unique products. $\overline{\text{Gap}}_{pre}$ and $\overline{\text{Gap}}_{post}$ denote the mean reporting gap (in value terms) in each group before and after the DRCM is implemented, Δ Means is the difference in means and DiD is the difference in the mean difference. The last column contains the average number of times an observation in observed over the 24 months window.

TABLE 3.A.3: DRCM INSTANCES IN IMPORTING COUNTRIES (CONTINUED)

Case	Grp	N	# prod	$\overline{\text{Gap}}_{pre}$	$\overline{\text{Gap}}_{post}$	Δ Means	DiD	Mean Freq.
28: HU, 7.2012	0	695480	7988	0.12	0.11	0.015		12
	1	990	24	0.76	0.11	0.65	-0.63	7
29: HU, 1.2015	0	775653	7502	0.11	0.097	0.017		13
	1	16381	178	0.089	0.04	0.048	-0.032	13
30: IE, 5.2011	0	362648	8209	-0.01	0.011	-0.021		11
	1	120	8	-0.35	-0.63	0.28	-0.3	12
31: IT, 1.2011	0	1070555	8930	0.037	0.018	0.019		14
	1	3986	17	0.33	0.045	0.29	-0.27	16
32: IT, 5.2016	0	1276117	9320	0.032	0.053	-0.02		14
	1	838	2	0.71	-0.032	0.75	-0.77	17
33: LT, 1.2008	0	499663	7144	-0.16	-0.051	-0.1		11
	1	733	39	-0.024	0.067	-0.092	-0.013	7.1
34: LT, 8.2019	0	501659	7316	-0.15	-0.16	0.011		9.2
	1	892	4	-0.051	-0.14	0.086	-0.075	12
35: LV, 10.2011	0	503982	7525	0.0067	-0.02	0.027		11
	1	357	16	-0.19	-0.41	0.22	-0.19	8.5
36: LV, 1.2016	0	587466	7043	0.019	-0.026	0.044		12
	1	318	7	-0.12	-0.29	0.17	-0.13	8.8
37: LV, 4.2016	0	591595	7387	0.013	-0.026	0.039		12
	1	2547	24	-0.51	-0.2	-0.31	0.35	12
38: LV, 7.2016	0	599854	7420	0.00055	-0.028	0.028		12
	1	744	23	-0.095	-0.2	0.1	-0.075	7.7
39: LV, 1.2017	0	613128	7458	-0.025	-0.033	0.0082		12
	1	574	16	-0.39	0.24	-0.63	0.63	9.3
40: LV, 1.2018	0	604893	6784	-0.034	-0.017	-0.017		12
	1	35136	404	-0.021	-0.052	0.031	-0.048	12
41: NL, 1.2007	0	763324	9379	0.42	0.41	0.0069		12
	1	2394	40	-0.11	0.046	-0.16	0.16	10
42: NL, 4.2013	0	822139	8344	0.36	0.39	-0.033		13
	1	4543	19	0.16	-0.0064	0.17	-0.2	15
43: PL, 7.2011	0	927517	8933	-0.046	0.053	-0.098		13
	1	3514	81	0.1	0.21	-0.11	0.0072	9.5
44: PL, 10.2013	0	967363	8108	0.059	0.069	-0.01		14
	1	24018	229	-0.0028	-0.048	0.045	-0.055	13
45: PL, 7.2015	0	1042105	8275	0.098	0.15	-0.055		14
	1	3811	28	0.54	0.24	0.3	-0.36	13
46: PL, 1.2017	0	1076376	8734	0.17	0.17	-0.0042		14
	1	3185	31	0.029	0.0072	0.022	-0.026	12
47: PL, 1.2018	0	1113873	8391	0.17	0.17	0.0035		14
	1	552	1	-0.56	-0.26	-0.3	0.31	21
48: PT, 10.2006	0	607266	9839	0.022	-0.042	0.064		12
	1	681	28	0.17	0.13	0.035	0.029	11
49: RO, 6.2011	0	777280	8668	-0.12	-0.1	-0.014		12
	1	975	22	-0.11	-0.1	-0.0091	-0.0048	7
50: RO, 1.2016	0	909676	8081	-0.073	-0.067	-0.0064		13
	1	4155	19	-0.29	-0.26	-0.034	0.028	15
51: SE, 1.2013	0	763087	7802	0.11	0.12	-0.01		14
	1	1000	26	-0.33	-0.24	-0.096	0.085	11
52: SI, 1.2010	0	550878	7950	-0.091	-0.06	-0.031		12
	1	929	30	-0.1	-0.068	-0.036	0.0052	9
53: SK, 4.2009	0	595701	8015	-0.0061	-0.017	0.011		12
	1	573	26	0.029	-0.23	0.26	-0.25	7.1
54: SK, 1.2014	0	647612	7500	0.076	0.1	-0.023		13
	1	29232	435	0.046	0.029	0.018	-0.041	12

Note: Grp (0/1) refers to treatment and control groups, i.e. those trade flows subject to the DRCM (1) or not (0). N is the number of observations in each group, and # prod refers to the number of unique products. $\overline{\text{Gap}}_{pre}$ and $\overline{\text{Gap}}_{post}$ denote the mean reporting gap (in value terms) in each group before and after the DRCM is implemented, Δ Means is the difference in means and DiD is the difference in the mean difference. The last column contains the average number of times an observation in observed over the 24 months window.

TABLE 3.A.4: DRCM INSTANCES IN EXPORTING COUNTRIES (CONTINUES IN NEXT TABLE)

Case	Grp	N	# prod	$\overline{\text{Gap}}_{pre}$	$\overline{\text{Gap}}_{post}$	Δ Means	DiD	Mean Freq.
1: AT, 7.2007	0	937679	8175	0.029	0.039	-0.01		11
	1	2736	56	-0.0076	-0.15	0.14	-0.15	12
2: AT, 1.2012	0	1007803	7920	0.098	0.081	0.017		13
	1	3550	17	-0.18	-0.083	-0.099	0.12	12
3: AT, 1.2014	0	1041732	7111	0.052	0.073	-0.021		13
	1	50249	434	-0.063	-0.026	-0.037	0.016	13
4: BG, 1.2014	0	155269	4574	0.32	0.33	-0.011		8.3
	1	1287	29	0.22	0.16	0.064	-0.075	9.7
5: CZ, 4.2011	0	773658	7990	0.11	0.14	-0.029		12
	1	3188	64	0.045	0.058	-0.013	-0.016	12
6: CZ, 4.2015	0	922503	7151	0.2	0.24	-0.038		13
	1	60717	540	0.075	0.08	-0.0056	-0.033	12
7: CZ, 7.2015	0	1008631	7704	0.2	0.24	-0.035		13
	1	1142	18	0.0026	0.17	-0.17	0.13	9.3
8: CZ, 9.2015	0	1026165	7717	0.21	0.24	-0.028		13
	1	4	1	NA	0.17	NA	NA	4
9: DE, 1.2011	0	2731608	8671	-0.036	-0.02	-0.016		17
	1	8257	61	0.066	-0.039	0.11	-0.12	14
10: DE, 7.2011	0	2732541	9392	-0.026	-0.028	0.002		16
	1	9461	17	0.063	-0.035	0.097	-0.095	22
11: DE, 10.2014	0	2899504	8483	-0.034	7.1e-05	-0.034		18
	1	18316	78	0.055	0.16	-0.1	0.068	14
12: DE, 1.2015	0	2922175	8487	-0.028	0.0098	-0.038		18
	1	2682	32	-0.13	-0.28	0.15	-0.18	9.2
13: DK, 7.2012	0	729781	6988	0.097	0.15	-0.049		11
	1	1220	28	-1.2	-1.2	0.06	-0.11	11
14: DK, 7.2014	0	787052	6566	0.15	0.17	-0.02		12
	1	4129	18	-0.0093	0.22	-0.23	0.21	14
15: EE, 1.2011	0	159072	4818	-0.044	0.0046	-0.049		10
	1	455	19	0.062	-0.18	0.24	-0.29	8.8
16: EE, 4.2012	0	172455	5059	-0.0047	-0.037	0.033		11
	1	57	3	0.68	1.1	-0.39	0.42	8.1
17: EE, 7.2014	0	190361	4818	-0.081	-0.071	-0.01		11
	1	170	9	-0.4	0.19	-0.59	0.58	6.3
18: EE, 1.2017	0	212338	5050	-0.071	-0.1	0.03		10
	1	4188	154	-0.072	0.023	-0.095	0.12	11
19: ES, 1.2004	0	733124	8780	0.079	0.023	0.056		11
	1	2973	64	0.079	0.23	-0.15	0.21	11
20: ES, 4.2015	0	1195640	8230	0.045	0.071	-0.026		13
	1	1875	11	0.44	0.18	0.26	-0.29	15
21: FI, 1.2015	0	400925	5111	-0.072	-0.063	-0.0086		11
	1	546	21	0.053	0.34	-0.28	0.28	10
22: FR, 1.2008	0	1669931	8737	0.06	0.052	0.0078		14
	1	3503	41	0.22	0.23	-0.013	0.021	15
23: GB, 6.2007	0	1326737	9377	-0.036	-0.071	0.035		12
	1	1195	3	0.26	0.29	-0.024	0.059	15
24: GR, 1.2007	0	145247	5389	-0.013	-0.082	0.069		6.7
	1	302	21	0.16	-0.055	0.22	-0.15	5.9
25: GR, 8.2017	0	253521	5816	0.15	0.2	-0.052		8.4
	1	871	4	0.31	0.14	0.18	-0.23	13
26: HR, 1.2019	0	220587	5395	0.23	0.26	-0.022		9.5
	1	189	12	0.3	0.31	-0.011	-0.011	7.9
27: HU, 1.2008	0	362446	4923	-0.18	-0.16	-0.024		11
	1	1462	35	-0.14	-0.094	-0.043	0.019	9.6

Note: Grp (0/1) refers to treatment and control groups, i.e. those trade flows subject to the DRCM (1) or not (0). N is the number of observations in each group, and # prod refers to the number of unique products. $\overline{\text{Gap}}_{pre}$ and $\overline{\text{Gap}}_{post}$ denote the mean reporting gap (in value terms) in each group before and after the DRCM is implemented, Δ Means is the difference in means and DiD is the difference in the mean difference. The last column contains the average number of times an observation in observed over the 24 months window.

TABLE 3.A.5: DRCM INSTANCES IN EXPORTING COUNTRIES (CONTINUED)

Case	Grp	N	# prod	$\overline{\text{Gap}}_{pre}$	$\overline{\text{Gap}}_{post}$	Δ Means	DiD	Mean Freq.
28: HU, 7.2012	0	481975	6451	-0.062	-0.035	-0.026		11
	1	1885	25	-0.14	0.0022	-0.14	0.12	9.2
29: HU, 1.2015	0	596494	6233	0.00019	0.032	-0.032		12
	1	9960	152	-0.082	-0.032	-0.05	0.018	11
30: IE, 5.2011	0	180593	5139	-0.0037	-0.042	0.038		10
	1	537	21	0.45	0.28	0.17	-0.13	14
31: IT, 1.2011	0	1787212	8512	0.076	0.096	-0.019		14
	1	4154	17	-0.068	-0.051	-0.017	-0.0025	13
32: IT, 5.2016	0	2047668	8891	0.15	0.15	-0.0031		14
	1	1055	2	0.88	0.51	0.37	-0.37	20
33: LT, 1.2008	0	163573	5419	0.15	0.054	0.097		9.5
	1	2092	42	0.18	0.049	0.14	-0.038	9.6
34: LT, 8.2019	0	252007	5774	0.15	0.19	-0.044		8.2
	1	607	4	-0.68	-0.17	-0.51	0.47	8.8
35: LV, 10.2011	0	167791	5645	-0.014	-0.0032	-0.01		10
	1	396	17	0.36	-0.02	0.38	-0.39	8.1
36: LV, 1.2016	0	206310	5364	0.052	0.058	-0.0062		11
	1	771	7	0.19	0.55	-0.35	0.35	13
37: LV, 4.2016	0	210075	5652	0.049	0.051	-0.0019		10
	1	972	14	0.82	0.74	0.086	-0.088	12
38: LV, 7.2016	0	215224	5717	0.06	0.045	0.015		10
	1	604	17	-0.26	-0.14	-0.12	0.13	7.4
39: LV, 1.2017	0	225231	5818	0.059	0.047	0.012		10
	1	156	8	0.69	0.49	0.2	-0.19	8.2
40: LV, 1.2018	0	228664	5320	0.048	0.048	-0.00069		11
	1	13271	336	0.035	-0.075	0.11	-0.11	10
41: NL, 1.2007	0	1483943	9269	-0.17	-0.17	-0.0038		13
	1	1998	38	-0.27	-0.36	0.093	-0.096	11
42: NL, 4.2013	0	1606328	8290	-0.11	-0.14	0.023		14
	1	9269	19	0.34	0.27	0.071	-0.048	19
43: PL, 7.2011	0	882182	7598	0.061	-0.024	0.085		12
	1	2692	64	0.11	0.13	-0.021	0.11	8.9
44: PL, 10.2013	0	1021302	6982	0.0033	0.027	-0.024		13
	1	22162	207	0.11	0.12	-0.012	-0.012	11
45: PL, 7.2015	0	1280712	7315	0.069	0.072	-0.0027		14
	1	5356	29	0.23	0.19	0.036	-0.039	14
46: PL, 1.2017	0	1416733	7740	0.06	0.054	0.0054		14
	1	3587	25	0.31	0.36	-0.051	0.056	12
47: PL, 1.2018	0	1477856	7452	0.055	0.097	-0.042		15
	1	640	1	0.67	0.51	0.17	-0.21	24
48: PT, 10.2006	0	224743	6312	-0.00055	0.053	-0.053		9.1
	1	824	34	0.16	0.31	-0.15	0.096	13
49: RO, 6.2011	0	203048	5387	0.039	0.086	-0.046		8.5
	1	968	17	-0.057	0.31	-0.37	0.32	7.2
50: RO, 1.2016	0	324986	5590	0.059	0.11	-0.047		9.7
	1	1604	18	0.091	0.021	0.07	-0.12	9.8
51: SE, 1.2013	0	831738	6741	-0.067	-0.086	0.019		13
	1	1087	28	0.16	0.036	0.13	-0.11	10
52: SI, 1.2010	0	210777	4666	0.088	0.071	0.017		8.7
	1	1217	37	0.23	0.2	0.027	-0.0093	11
53: SK, 4.2009	0	285023	5663	0.058	0.074	-0.016		10
	1	1129	30	0.079	-0.015	0.095	-0.11	10
54: SK, 1.2014	0	367479	5546	0.02	0.034	-0.015		12
	1	18957	303	0.16	0.17	-0.011	-0.0035	12

Note: Grp (0/1) refers to treatment and control groups, i.e. those trade flows subject to the DRCM (1) or not (0). N is the number of observations in each group, and # prod refers to the number of unique products. $\overline{\text{Gap}}_{pre}$ and $\overline{\text{Gap}}_{post}$ denote the mean reporting gap (in value terms) in each group before and after the DRCM is implemented, Δ Means is the difference in means and DiD is the difference in the mean difference. The last column contains the average number of times an observation in observed over the 24 months window.

Table 3.A.6: Static difference-in-difference: value gap — Balanced sample

	Importer sample				
	Dependant variable: Value gap, log				
	(1)	(2)	(3)	(4)	(5)
DRCM (D_{jt})	-0.0415***	-0.0367***	-0.0388***	-0.0373***	-0.0413***
	(-3.89)	(-2.70)	(-3.63)	(-3.52)	(-2.92)
Adjusted R^2	0.486	0.492	0.487	0.487	0.522
Within R^2	0.000	0.000	0.000	0.000	0.000
Observations	24,394,660	24,216,387	24,394,606	24,393,661	19,701,142
	Exporter sample				
	Dependant variable: Value gap, log				
	(1)	(2)	(3)	(4)	(5)
DRCM (D_{jt})	-0.0211**	-0.00816	-0.0196*	-0.0195*	-0.0268
	(-2.02)	(-0.54)	(-1.83)	(-1.81)	(-1.53)
Adjusted R^2	0.493	0.503	0.494	0.494	0.519
Within R^2	0.000	0.000	0.000	0.000	0.000
Observations	24,764,482	24,599,503	24,764,466	24,763,918	17,519,337
t, j FE	✓	✓	✓	✓	✓
ymp FE		✓			
ymc FE			✓		✓
ymk FE			✓		
ymck FE				✓	
ymkp FE					✓

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Fixed effects (FE): event time (t), unit ($j \equiv iepv$) calendar time (ym), product (p), country in which the reform takes place ($c = i$ in the importer sample, and $c = e$ in the exporting sample), partner country ($k = e$ in the importer sample, and $k = i$ in the exporting sample). The dependent variable is the reporting gap in terms of value, in logs as specified in expression (3.1). Sample balanced such that each unit j appears at least 20 times in the 24-month window.

Table 3.A.7: Static difference-in-difference: value gap — No event overlaps

	Importer sample				
	Dependant variable: Value gap, log				
	(1)	(2)	(3)	(4)	(5)
DRCM (D_{jt})	-0.0481***	-0.0443***	-0.0435***	-0.0420***	-0.0532***
	(-3.62)	(-2.67)	(-3.25)	(-3.15)	(-2.73)
Adjusted R^2	0.471	0.476	0.472	0.472	0.501
Within R^2	0.000	0.000	0.000	0.000	0.000
Observations	30,913,101	30,747,517	30,913,082	30,912,835	24,756,295
	Exporter sample				
	Dependant variable: Value gap, log				
	(1)	(2)	(3)	(4)	(5)
DRCM (D_{jt})	-0.00758	0.00349	-0.00795	-0.00817	-0.00514
	(-0.70)	(0.21)	(-0.73)	(-0.74)	(-0.25)
Adjusted R^2	0.480	0.487	0.481	0.482	0.495
Within R^2	0.000	0.000	0.000	0.000	0.000
Observations	29395610	29220157	29395610	29395398	20576550
t, j FE	✓	✓	✓	✓	✓
ymp FE		✓			
ymc FE			✓		✓
ymk FE			✓		
$ymck$ FE				✓	
$ymkp$ FE					✓

Note: *p$<$0.1; **p$<$0.05; ***p$<$0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Fixed effects (FE): event time (t), unit ($j \equiv iepv$) calendar time (ym), product (p), country in which the reform takes place ($c = i$ in the importer sample, and $c = e$ in the exporting sample), partner country ($k = e$ in the importer sample, and $k = i$ in the exporting sample). The dependent variable is the reporting gap in terms of value, in logs as specified in expression (3.1). Overlaps in events are dropped from the sample.

Table 3.A.8: Static difference-in-difference: value gap — Without outliers

	Importer sample				
	Dependant variable: Value gap, log				
	(1)	(2)	(3)	(4)	(5)
DRCM (D_{jt})	-0.0312***	-0.0182*	-0.0281***	-0.0266***	-0.0229**
	(-3.68)	(-1.77)	(-3.25)	(-3.09)	(-2.10)
Adjusted R^2	0.444	0.448	0.445	0.445	0.475
Within R^2	0.000	0.000	0.000	0.000	0.000
Observations	35,379,360	35,219,033	35,379,342	35,379,061	29,251,128
	Exporter sample				
	Dependant variable: Value gap, log				
	(1)	(2)	(3)	(4)	(5)
DRCM (D_{jt})	-0.0139*	-0.00413	-0.0142*	-0.0142*	-0.0167
	(-1.84)	(-0.36)	(-1.84)	(-1.83)	(-1.24)
Adjusted R^2	0.448	0.454	0.449	0.449	0.467
Within R^2	0.000	0.000	0.000	0.000	0.000
Observations	35,733,394	35,568,587	35,733,394	35,733,154	26,580,851
t, j FE	✓	✓	✓	✓	✓
ymp FE		✓			
ymc FE			✓		✓
ymk FE			✓		
$ymck$ FE				✓	
$ymkp$ FE					✓

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Fixed effects (FE): event time (t), unit ($j \equiv iepv$) calendar time (ym), product (p), country in which the reform takes place ($c = i$ in the importer sample, and $c = e$ in the exporting sample), partner country ($k = e$ in the importer sample, and $k = i$ in the exporting sample). The dependent variable is the reporting gap in terms of value, in logs as specified in expression (3.1). The top and bottom percentiles in terms of the reporting gaps are dropped from the sample.

TABLE 3.A.9: STATIC DIFFERENCE-IN-DIFFERENCE: VALUE GAP — ADDITIONAL CONTROLS

	Importer sample			
	Dependant variable: Value gap, log			
	(1)	(2)	(3)	(4)
DRCM (D_{jt})	-0.0323***	-0.0343***	-0.0205*	-0.0235*
	(-3.20)	(-3.42)	(-1.68)	(-1.95)
Importer VAT main rate	0.00648		0.00686	
	(1.45)		(1.53)	
Exporter VAT main rate	-0.000690		-0.000396	
	(-0.25)		(-0.14)	
Importer Intrastat threshold	0.0426**		0.0385*	
	(2.00)		(1.81)	
Exporter Intrastat threshold	-0.0564***		-0.0543***	
	(-4.11)		(-3.99)	
Adjusted R^2	0.480	0.480	0.484	0.484
Within R^2	0.000	0.000	0.000	0.000
Observations	35,413,703	35,413,703	35,269,247	35,269,247
	Exporter sample			
	Dependant variable: Value gap, log			
	(1)	(2)	(3)	(4)
DRCM (D_{jt})	-0.0104	-0.0111	0.00181	0.00219
	(-1.18)	(-1.26)	(0.14)	(0.17)
Importer VAT main rate	-0.00185		-0.00156	
	(-0.85)		(-0.71)	
Exporter VAT main rate	-0.00803*		-0.00940**	
	(-1.71)		(-1.97)	
Importer Intrastat threshold	0.109***		0.110***	
	(4.74)		(4.71)	
Exporter Intrastat threshold	-0.0528		-0.0568	
	(-1.58)		(-1.62)	
Adjusted R^2	0.484	0.484	0.490	0.490
Within R^2	0.000	0.000	0.000	0.000
Observations	36,192,505	36,192,505	36,045,319	36,045,319
t, j FE	✓	✓	✓	✓
ymp FE			✓	✓

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Fixed effects (FE): event time (t), unit ($j \equiv iepv$) calendar time (ym), product (p), country in which the reform takes place ($c = i$ in the importer sample, and $c = e$ in the exporting sample), partner country ($k = e$ in the importer sample, and $k = i$ in the exporting sample). The dependent variable is the reporting gap in terms of value, in logs as specified in expression (3.1).

Table 3.A.10: Static difference-in-difference: value gap — Reforms in partner country

	Importer sample				
	Dependant variable: Value gap, log				
	(1)	(2)	(3)	(4)	(5)
DRCM (D_{jt})	-0.0349***	-0.0245**	-0.0314***	-0.0302***	-0.0280**
	(-3.51)	(-2.04)	(-3.15)	(-3.04)	(-2.22)
DRCM partner	0.0127	0.00525	-0.00611	-0.00738	
	(0.92)	(0.37)	(-0.41)	(-0.51)	
Adjusted R^2	0.478	0.483	0.479	0.479	0.508
Within R^2	0.000	0.000	0.000	0.000	0.000
Observations	36,115,872	35,957,790	36,115,854	36,115,584	29,952,923
	Exporter sample				
	Dependant variable: Value gap, log				
	(1)	(2)	(3)	(4)	(5)
DRCM (D_{jt})	-0.0101	0.00184	-0.0101	-0.00987	-0.0102
	(-1.14)	(0.14)	(-1.14)	(-1.10)	(-0.66)
DRCM partner	-0.0132	-0.0132	-0.00776	-0.00850	
	(-1.12)	(-1.02)	(-0.70)	(-0.76)	
Adjusted R^2	0.484	0.491	0.485	0.485	0.502
Within R^2	0.000	0.000	0.000	0.000	0.000
Observations	36,477,528	36,314,706	36,477,528	36,477,290	27,287,522
t, j FE	✓	✓	✓	✓	✓
ymp FE		✓			
ymc FE			✓		✓
ymk FE			✓		
ymck FE				✓	
ymkp FE					✓

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Fixed effects (FE): event time (t), unit ($j \equiv iepv$) calendar time (ym), product (p), country in which the reform takes place ($c = i$ in the importer sample, and $c = e$ in the exporting sample), partner country ($k = e$ in the importer sample, and $k = i$ in the exporting sample). The dependent variable is the reporting gap in terms of value, in logs as specified in expression (3.1). *DRCM partner* cannot be included in the last column, as it would be absorbed by the *ymkp* fixed effect.

TABLE 3.A.11: BASELINE STATIC DIFFERENCE-IN-DIFFERENCE: QUANTITY GAP

	Importer sample				
	Dependant variable: Quantity gap, log				
	(1)	(2)	(3)	(4)	(5)
DRCM (D_{jt})	-0.0341***	-0.0328**	-0.0323***	-0.0299***	-0.0376***
	(-2.98)	(-2.40)	(-2.83)	(-2.64)	(-2.61)
Adjusted R^2	0.501	0.503	0.501	0.501	0.524
Within R^2	0.000	0.000	0.000	0.000	0.000
Observations	32,137,670	31,960,693	32,137,645	32,137,357	26,617,025
	Exporter sample				
	Dependant variable: Quantity gap, log				
	(1)	(2)	(3)	(4)	(5)
DRCM (D_{jt})	-0.00875	-0.00590	-0.00818	-0.00800	-0.0145
	(-0.81)	(-0.39)	(-0.77)	(-0.74)	(-0.85)
Adjusted R^2	0.504	0.509	0.505	0.505	0.521
Within R^2	0.000	0.000	0.000	0.000	0.000
Observations	32,527,622	32,357,298	32,527,622	32,527,363	24,218,138
t, j FE	✓	✓	✓	✓	✓
ymp FE		✓			
ymc FE			✓		✓
ymk FE			✓		
ymck FE				✓	
ymkp FE					✓

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Fixed effects (FE): event time (t), unit ($j \equiv iepv$) calendar time (ym), product (p), country in which the reform takes place ($c = i$ in the importer sample, and $c = e$ in the exporting sample), partner country ($k = e$ in the importer sample, and $k = i$ in the exporting sample). The dependent variable is the reporting gap in terms of quantities, in logs as specified in expression (3.1).

Table 3.A.12: Static difference-in-difference: value gap — Restricted control group

	Importer sample				
	Dependant variable: Value gap, log				
	(1)	(2)	(3)	(4)	(5)
DRCM (D_{jt})	-0.0509***	-0.0655***	-0.0474***	-0.0454***	-0.0776***
	(-3.67)	(-2.99)	(-3.42)	(-3.30)	(-2.79)
Adjusted R^2	0.472	0.477	0.472	0.473	0.510
Within R^2	0.000	0.000	0.000	0.000	0.000
Observations	7,135,345	7,064,311	7,135,291	7,134,556	4,815,918
	Exporter sample				
	Dependant variable: Value gap, log				
	(1)	(2)	(3)	(4)	(5)
DRCM (D_{jt})	-0.00423	0.0232	-0.00379	-0.00199	-0.00257
	(-0.38)	(0.90)	(-0.31)	(-0.16)	(-0.09)
Adjusted R^2	0.474	0.482	0.475	0.476	0.498
Within R^2	0.000	0.000	0.000	0.000	0.000
Observations	6,988,820	6,916,054	6,988,806	6,988,055	4,055,657
t, j FE	✓	✓	✓	✓	✓
ymp FE		✓			
ymc FE			✓		✓
ymk FE			✓		
ymck FE				✓	
ymkp FE					✓

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Fixed effects (FE): event time (t), unit ($j \equiv iepv$) calendar time (ym), product (p), country in which the reform takes place ($c = i$ in the importer sample, and $c = e$ in the exporting sample), partner country ($k = e$ in the importer sample, and $k = i$ in the exporting sample). The dependent variable is the reporting gap in terms of value, in logs as specified in expression (3.1). The sample is restricted to goods belonging to the same CN section as the treated units.

TABLE 3.A.13: STATIC DIFFERENCE-IN-DIFFERENCE: VALUE GAP — INTERACTIONS

	Importer sample				
	Dependant variable: Value gap, log				
	(1)	(2)	(3)	(4)	(5)
DRCM (D_{jt})	0.176**	0.223**	0.0824	0.0706	-0.0365
	(1.96)	(2.21)	(0.93)	(0.80)	(-0.36)
Importer VAT main rate	0.00649	0.00686			
	(1.45)	(1.53)			
DRCM (D_{jt}) $\times$ Imp. VAT rate	-0.00990**	-0.0114**	-0.00538	-0.00475	0.000431
	(-2.38)	(-2.46)	(-1.31)	(-1.17)	(0.09)
Exporter VAT main rate	-0.000693	-0.000400			
	(-0.25)	(-0.14)			
Importer Intrastat threshold	0.0429**	0.0389*			
	(2.02)	(1.83)			
Exporter Intrastat threshold	-0.0564***	-0.0542***			
	(-4.11)	(-3.99)			
Adjusted R^2	0.480	0.484	0.481	0.481	0.509
Within R^2	0.000	0.000	0.000	0.000	0.000
Observations	35,413,703	35,269,247	35,413,692	35,413,430	29,602,215
	Exporter sample				
	Dependant variable: Value gap, log				
	(1)	(2)	(3)	(4)	(5)
DRCM (D_{jt})	-0.154	-0.180	-0.114	-0.141	-0.0576
	(-1.64)	(-1.49)	(-1.22)	(-1.49)	(-0.38)
Importer VAT main rate	-0.00803*	-0.00940**			
	(-1.71)	(-1.97)			
DRCM (D_{jt}) $\times$ Imp. VAT rate	0.00686	0.00854	0.00491	0.00624	0.00221
	(1.53)	(1.52)	(1.10)	(1.38)	(0.31)
Exporter VAT main rate	-0.00185	-0.00156			
	(-0.85)	(-0.71)			
Importer Intrastat threshold	0.109***	0.110***			
	(4.74)	(4.71)			
Exporter Intrastat threshold	-0.0531	-0.0572			
	(-1.59)	(-1.63)			
Adjusted R^2	0.484	0.490	0.484	0.484	0.502
Within R^2	0.000	0.000	0.000	0.000	0.000
Observations	36,192,505	36,045,319	36,192,505	36,192,288	27,238,333
t, j FE	✓	✓	✓	✓	✓
ymp FE		✓			
ymc FE			✓		✓
ymk FE			✓		
ymck FE				✓	
ymkp FE					✓

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Fixed effects (FE): event time (t), unit ($j \equiv iepv$) calendar time (ym), product (p), country in which the reform takes place ($c = i$ in the importer sample, and $c = e$ in the exporting sample), partner country ($k = e$ in the importer sample, and $k = i$ in the exporting sample). The dependent variable is the reporting gap in terms of value, in logs as specified in expression (3.1).

TABLE 3.A.14: STATIC DIFF.-IN-DIFF.: VALUE GAP — FRAUD SHIFTING IN IMPORTER SAMPLE

	Unit value				
	Dependant variable: Value gap, log				
	(1)	(2)	(3)	(4)	(5)
DRCM (D_{jt})	-0.0482***	-0.0410**	-0.0436***	-0.0420***	-0.0508***
	(-3.63)	(-2.47)	(-3.26)	(-3.16)	(-2.62)
High UV × post	-0.00281	0.0250***	-0.00231	-0.00251	0.0149
	(-0.53)	(2.65)	(-0.44)	(-0.48)	(1.40)
Adjusted R^2	0.471	0.476	0.472	0.472	0.501
Within R^2	0.000	0.000	0.000	0.000	0.000
Observations	30,913,101	30,747,517	30,913,082	30,912,835	24,756,295
	Goods subjectable to DRCM				
	Dependant variable: Value gap, log				
	(1)	(2)	(3)	(4)	(5)
DRCM (D_{jt})	-0.0491***	-0.0452***	-0.0444***	-0.0428***	-0.0504**
	(-3.69)	(-2.75)	(-3.32)	(-3.21)	(-2.57)
DRCM-good × post	-0.0131***	-0.00127	-0.0125***	-0.0125***	0.00437
	(-3.70)	(-0.23)	(-3.55)	(-3.54)	(0.69)
Adjusted R^2	0.471	0.476	0.472	0.472	0.501
Within R^2	0.000	0.000	0.000	0.000	0.000
Observations	30,913,101	30,747,517	30,913,082	30,912,835	24,756,295
	Goods within same HS4				
	Dependant variable: Value gap, log				
	(1)	(2)	(3)	(4)	(5)
DRCM (D_{jt})	-0.0481***	-0.0440***	-0.0434***	-0.0419***	-0.0529***
	(-3.62)	(-2.65)	(-3.25)	(-3.14)	(-2.71)
Same HS4 × post	0.0199	0.0640***	0.0217	0.0224	0.0487**
	(1.00)	(2.62)	(1.11)	(1.14)	(2.08)
Adjusted R^2	0.471	0.476	0.472	0.472	0.501
Within R^2	0.000	0.000	0.000	0.000	0.000
Observations	30,913,101	30,747,517	30,913,082	30,912,835	24,756,295
t, j FE	✓	✓	✓	✓	✓
ymp FE		✓			
ymc FE			✓		✓
ymk FE			✓		
ymck FE				✓	
ymkp FE					✓

Note: *p<0.1; **p<0.05; ***p<0.01. t-stats shown in parenthesis. Standard errors are clustered at the country pair level. Fixed effects (FE): event time (t), unit ($j \equiv iepv$) calendar time (ym), product (p), country in which the reform takes place ($c = i$ in the importer sample, and $c = e$ in the exporting sample), partner country ($k = e$ in the importer sample, and $k = i$ in the exporting sample). The dependent variable is the reporting gap in terms of value, in logs as specified in expression (3.1). *DRCM-good × post* is a dummy that takes value 1 post-reform if the good has high unit value (in top 5 percentiles in terms of unit value over whole sample). *DRCM-good × post* is a dummy that takes value 1 post reform if the good can be subject to the DRCM according to the law, but has not been in a specific reform. *Same HS4 × post* is a dummy that takes value 1 post reform if the good is within the same HS heading as the treated product - but has not been subject to the DRCM.

TABLE 3.A.15: ESTIMATED AMOUNTS OF FRAUD PRE-REFORM

					Event-specific $\hat{\beta}_v$			Common $\hat{\beta}$		
v	Trade (m. EUR)	VAT rate	$\hat{\beta}$	$\hat{\beta}_v$	Fraud (m. EUR)	Fraud (% rev.)	Fraud (% VAT gap)	Fraud (m. EUR)	Fraud (% rev.)	Fraud (% VAT gap)
1	813.10	20.00	-0.03					4.88	0.02	0.14
2	1,169.39	20.00	-0.03					7.02	0.03	0.29
3	6,939.16	20.00	-0.03					41.63	0.16	1.62
4	184.78	20.00	-0.03					1.11	0.03	0.10
5	547.90	20.00	-0.03	-0.28	30.82	0.26	0.70	3.29	0.03	0.07
6	10,616.61	21.00	-0.03	-0.02	39.33	0.32	1.48	66.88	0.55	2.51
7	267.80	21.00	-0.03					1.69	0.01	0.06
8	27.78	21.00	-0.03					0.18	0.00	0.01
9	6,081.43	19.00	-0.03	-0.03	37.80	0.02	0.13	34.66	0.02	0.12
10	9,178.48	19.00	-0.03					52.32	0.03	0.19
11	15,093.22	19.00	-0.03	-0.05	138.99	0.07	0.52	86.03	0.04	0.32
12	1,393.84	19.00	-0.03					7.94	0.00	0.03
13	90.39	25.00	-0.03	-0.09	1.95	0.01	0.06	0.68	0.00	0.02
14	1,861.91	25.00	-0.03					13.96	0.06	0.46
15	14.66	20.00	-0.03					0.09	0.01	0.03
16	7.52	20.00	-0.03	-0.04	0.06	0.00	0.03	0.05	0.00	0.02
17	16.03	20.00	-0.03					0.10	0.01	0.05
18	295.94	20.00	-0.03	-0.05	2.93	0.14	2.44	1.78	0.08	1.48
19	1,584.97	16.00	-0.03	-0.03	6.86	0.01	0.21	7.61	0.01	0.23
20	3,697.87	21.00	-0.03	-0.15	116.24	0.17	3.62	23.30	0.03	0.73
21	34.86	24.00	-0.03	-0.05	0.45	0.00	0.04	0.25	0.00	0.02
22	1,663.28	19.60	-0.03					9.78	0.01	0.03
23	3,919.17	17.50	-0.03					20.58	0.01	0.09
24	102.51	19.00	-0.03	-0.22	4.36	0.02	0.05	0.58	0.00	0.01
25	501.44	24.00	-0.03	-0.30	35.99	0.25	0.50	3.61	0.03	0.05
27	54.39	20.00	-0.03	-0.09	0.95	0.01	0.03	0.33	0.00	0.01
28	561.23	27.00	-0.03	-0.45	67.96	0.73	2.59	4.55	0.05	0.17
29	1,248.09	27.00	-0.03	-0.00	0.27	0.00	0.01	10.11	0.09	0.49
30	6.70	21.00	-0.03					0.04	0.00	0.00
31	3,421.78	20.00	-0.03	-0.25	171.50	0.17	0.45	20.53	0.02	0.05
32	1,366.48	22.00	-0.03	-0.63	190.02	0.19	0.51	9.02	0.01	0.02
33	48.68	18.00	-0.03	-0.03	0.31	0.01	0.03	0.26	0.01	0.02
35	87.59	22.00	-0.03	-0.23	4.47	0.31	0.45	0.58	0.04	0.06
36	9.68	21.00	-0.03	-0.16	0.33	0.02	0.11	0.06	0.00	0.02
37	398.69	21.00	-0.03					2.51	0.12	0.81
38	120.48	21.00	-0.03	-0.03	0.80	0.04	0.26	0.76	0.04	0.25
39	25.38	21.00	-0.03					0.16	0.01	0.04
40	784.00	21.00	-0.03	-0.05	7.96	0.34		4.94	0.21	
41	1,591.64	19.00	-0.03					9.07	0.02	11.46
42	3,435.34	21.00	-0.03	-0.16	118.07	0.28	2.48	21.64	0.05	0.46
43	529.31	23.00	-0.03					3.65	0.01	0.06
44	4,761.31	23.00	-0.03	-0.05	58.18	0.21	0.57	32.85	0.12	0.32
45	5,027.55	23.00	-0.03	-0.36	412.40	1.37	4.32	34.69	0.12	0.36
46	1,161.34	23.00	-0.03	-0.24	62.94	0.18	1.11	8.01	0.02	0.14
47	267.51	23.00	-0.03					1.85	0.00	
48	173.27	21.00	-0.03	-0.09	3.30	0.02	0.34	1.09	0.01	0.11
49	750.71	24.00	-0.03	-0.16	28.35	0.24	0.26	5.41	0.05	0.05
50	1,507.42	20.00	-0.03	-0.02	4.78	0.04	0.08	9.04	0.08	0.15
51	391.12	25.00	-0.03					2.93	0.01	0.21
52	130.19	20.00	-0.03	-0.32	8.30	0.26	2.17	0.78	0.02	0.20
53	224.05	19.00	-0.03	-0.24	10.12	0.22	0.39	1.28	0.03	0.05
54	5,557.51	20.00	-0.03	-0.04	46.33	0.92	2.19	33.35	0.66	1.58

Note: v refers to the event number. Data: importer sample. Amounts of trade and fraud, when not expressed as fractions, are in 2015 million EUR. (% rev.) means expressed in percentage of VAT revenues in the year of reform. (% VAT gap) means expressed in percentage of the VAT gap in the year of reform. All amounts are annual values. Missing values in columns 5 to 8 arise when $\hat{\beta}_v > 0$.

Appendix 3.B Additional figures

FIGURE 3.B.1: ILLUSTRATION OF MTIC FRAUD

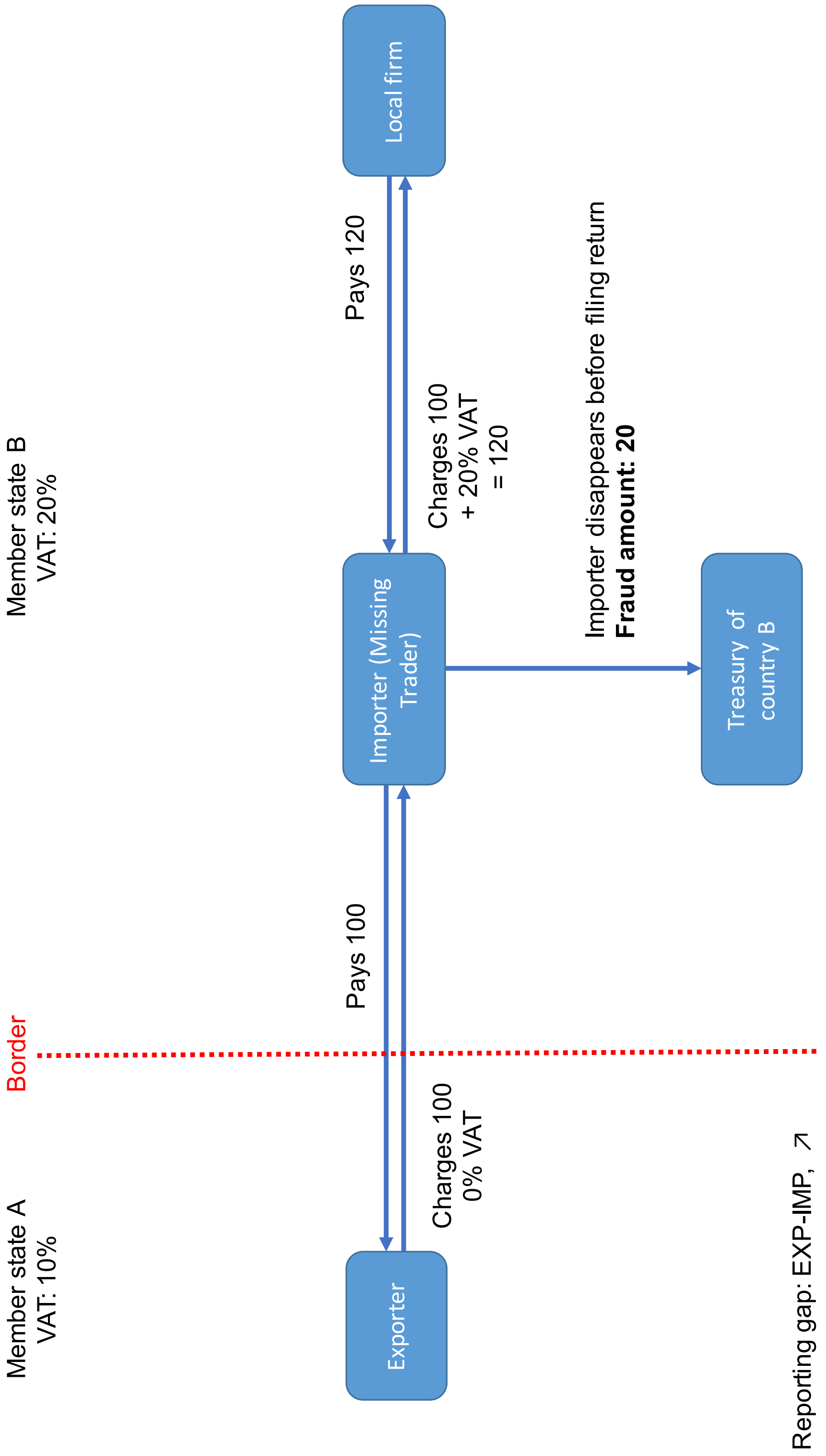


FIGURE 3.B.2: ILLUSTRATION OF IMPORTS DISGUISED AS DOMESTIC PURCHASES

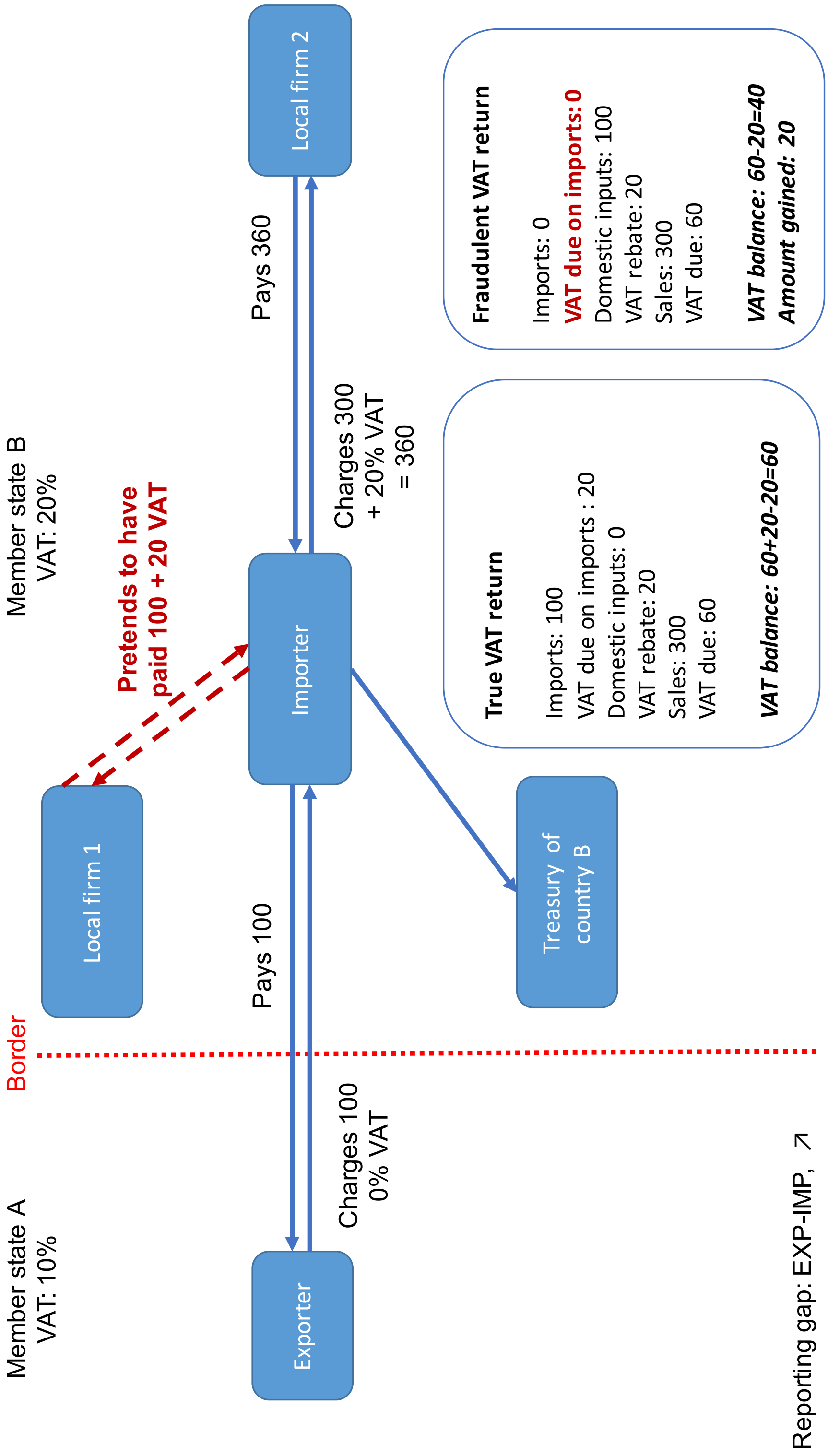


Figure 3.B.3: Illustration of domestic sales disguised as exports

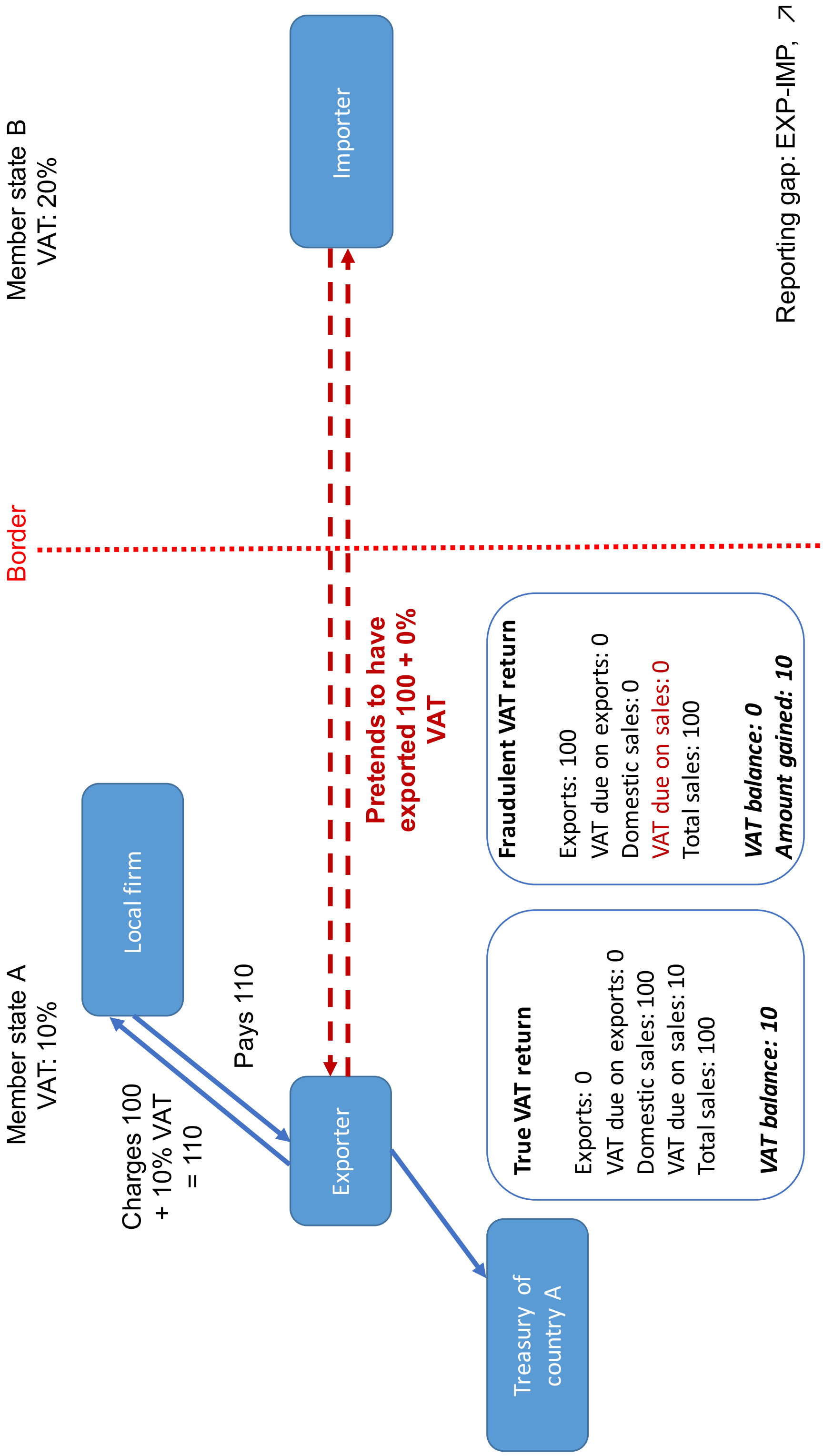
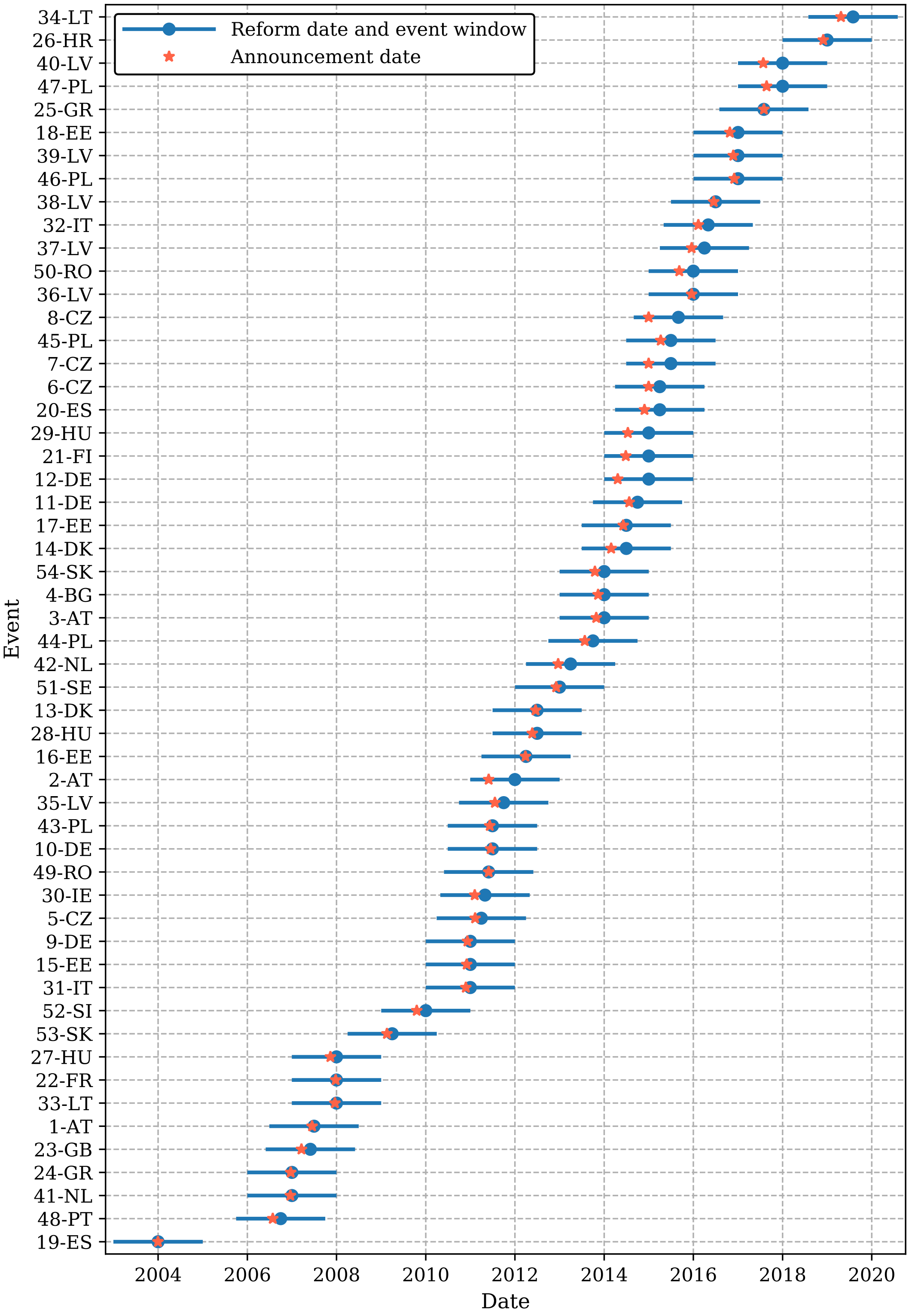
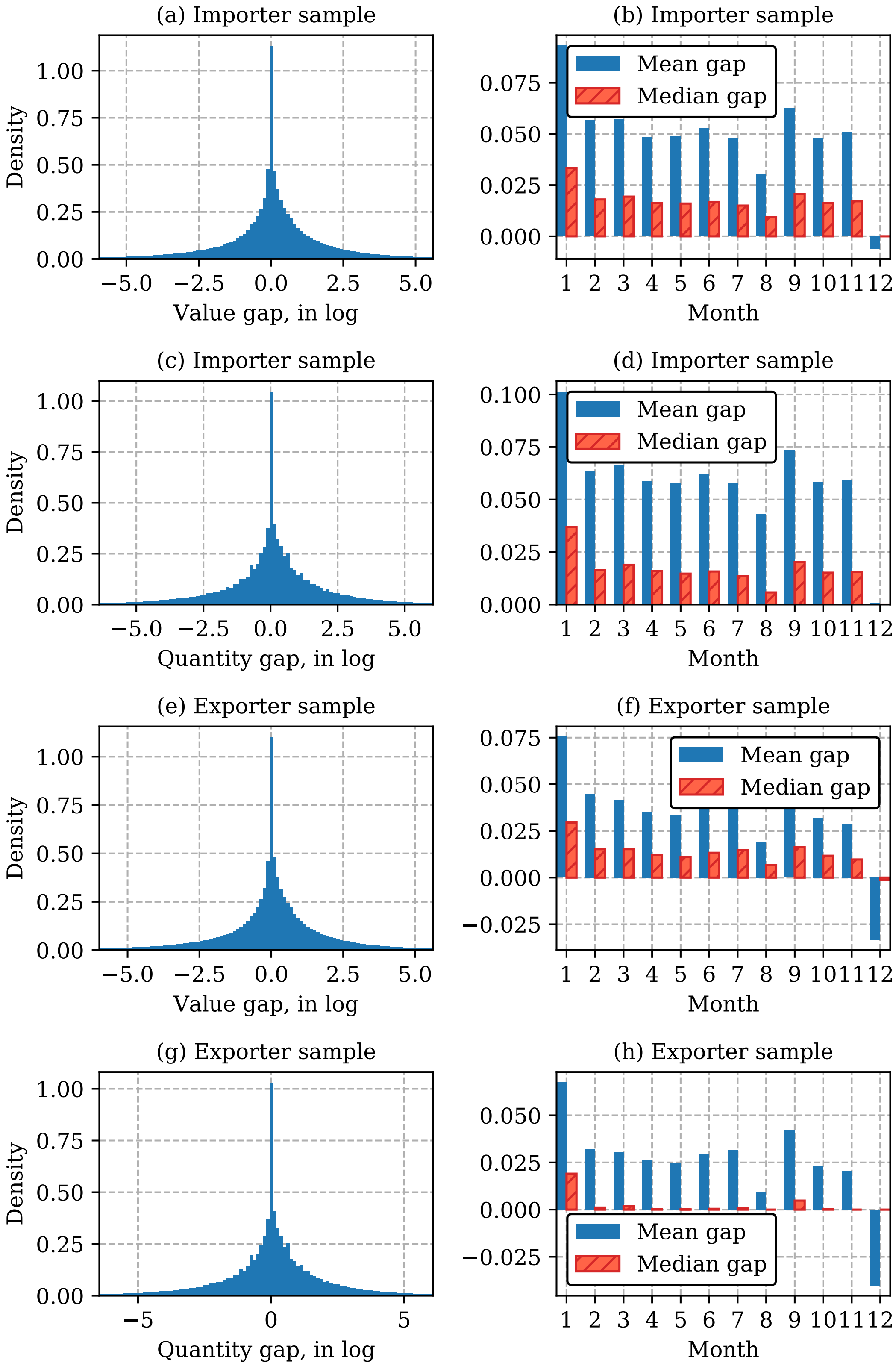


FIGURE 3.B.4: EVENT WINDOWS OVER TIME



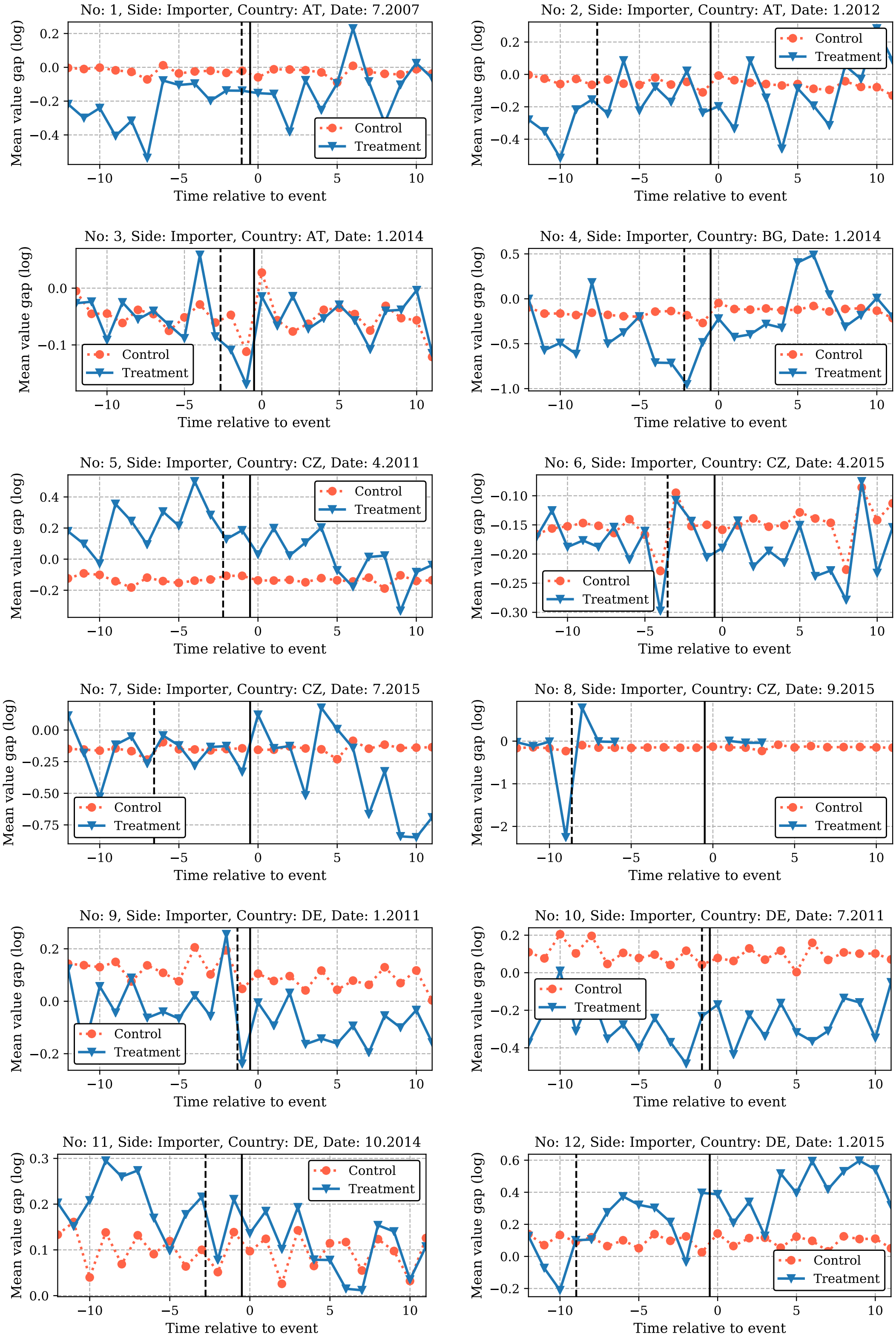
Note: Each line represents an event window of 12 months before and after the date of implementation of the DRCM, which is depicted as a round dot. The orange star depicts the date at which the reform is announced.

Figure 3.B.5: Histograms and monthly averages and medians of the reporting gaps



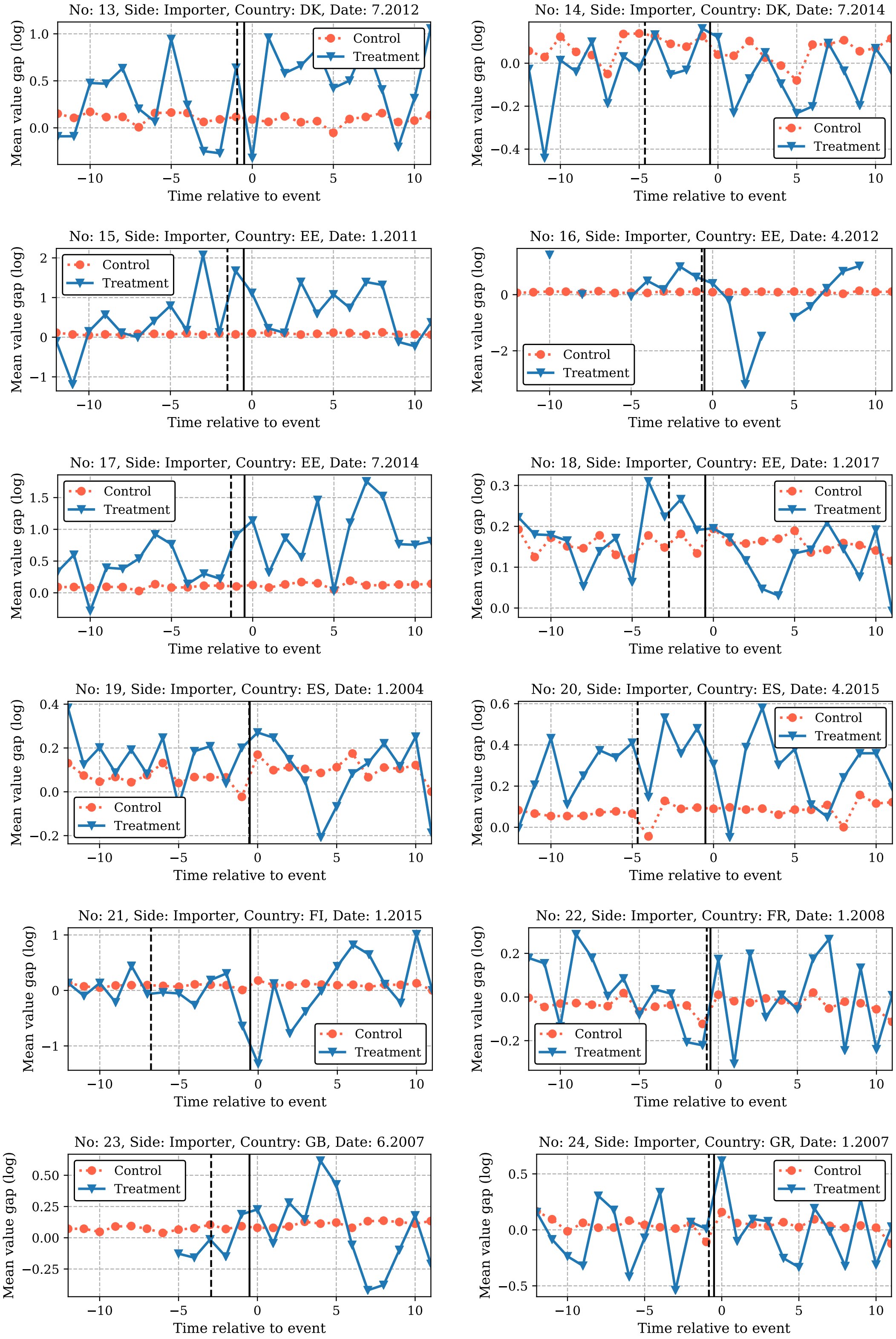
Note: Histograms and monthly averages of the reporting gaps in terms of values and quantities, for the importer and the exporter samples. The series underlying the figures in each row are the same. The bottom and top percentiles in terms of the reporting gaps are removed from the sample.

FIGURE 3.B.6: MEAN GAPS PER GROUP OVER TIME — EVENT IN IMPORTING COUNTRY (CONTINUES IN NEXT FIGURE)



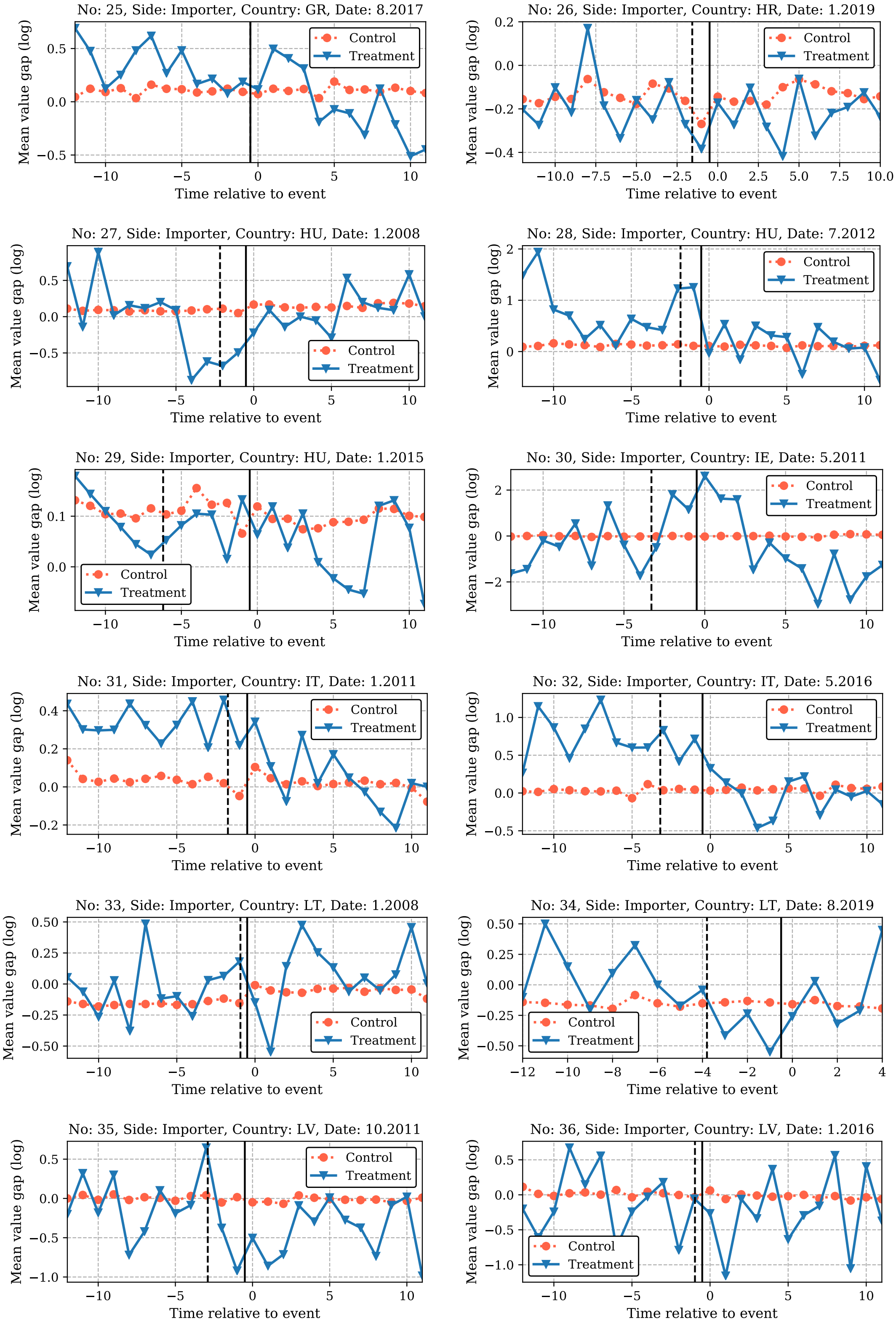
Note: Time series of the mean of gaps as calculated in (3.1), for the treatment group (i.e. products to which the DRCM applies at time 0) and the control group (the remaining products). The solid and dashed vertical lines indicate the date at which the DRCM starts to apply and the date at which the reform was announced.

FIGURE 3.B.7: MEAN GAPS PER GROUP OVER TIME — EVENT IN IMPORTING COUNTRY (CONTINUES IN NEXT FIGURE)



Note: Time series of the mean of gaps as calculated in (3.1), for the treatment group (i.e. products to which the DRCM applies at time 0) and the control group (the remaining products). The solid and dashed vertical lines indicate the date at which the DRCM starts to apply and the date at which the reform was announced.

Figure 3.B.8: Mean gaps per group over time — Event in importing country (continues in next figure)



No: 25, Side: Importer, Country: GR, Date: 8.2017

No: 26, Side: Importer, Country: HR, Date: 1.2019

No: 27, Side: Importer, Country: HU, Date: 1.2008

No: 28, Side: Importer, Country: HU, Date: 7.2012

No: 29, Side: Importer, Country: HU, Date: 1.2015

No: 30, Side: Importer, Country: IE, Date: 5.2011

No: 31, Side: Importer, Country: IT, Date: 1.2011

No: 32, Side: Importer, Country: IT, Date: 5.2016

No: 33, Side: Importer, Country: LT, Date: 1.2008

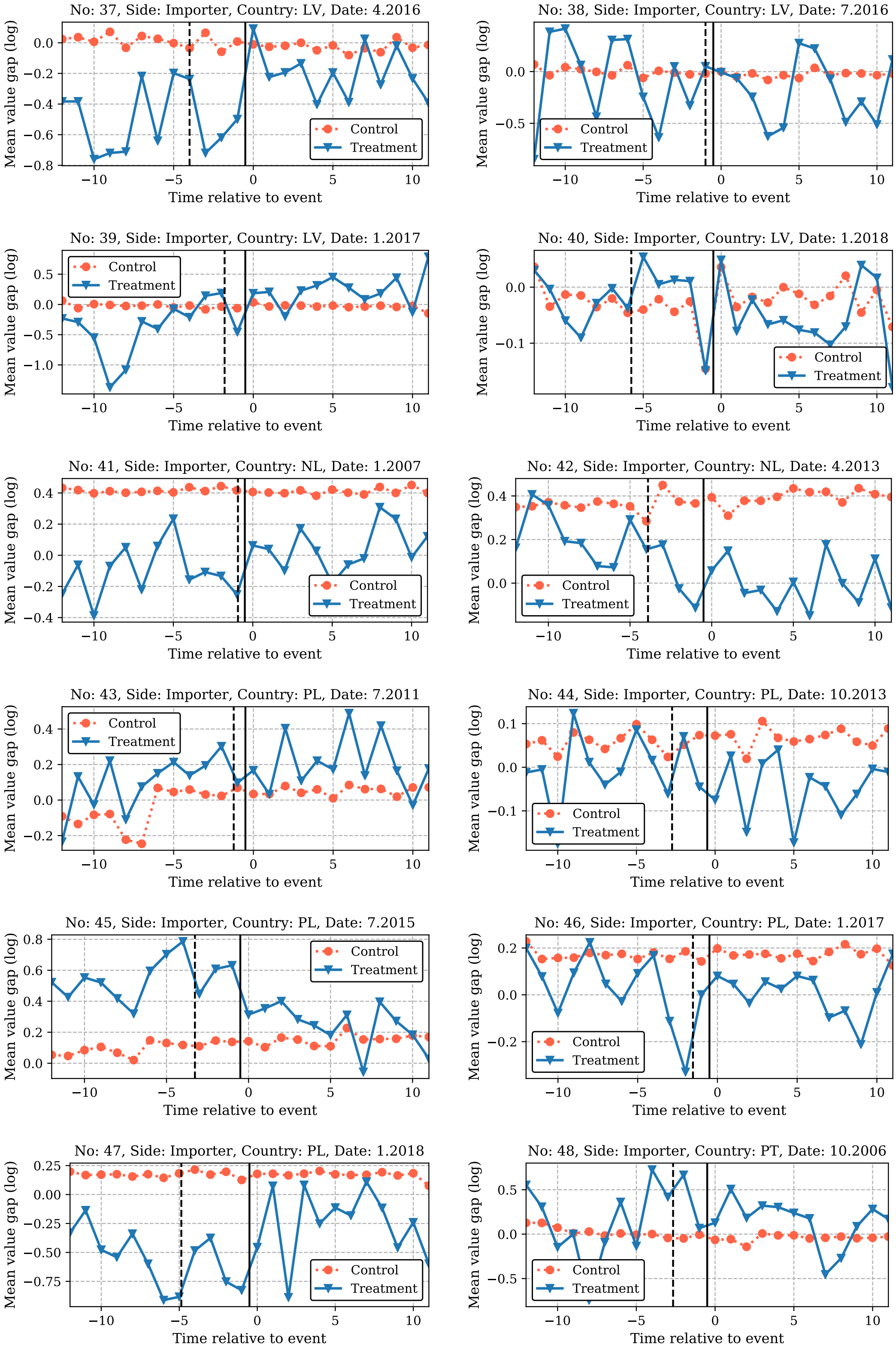
No: 34, Side: Importer, Country: LT, Date: 8.2019

No: 35, Side: Importer, Country: LV, Date: 10.2011

No: 36, Side: Importer, Country: LV, Date: 1.2016

Note: Time series of the mean of gaps as calculated in (3.1), for the treatment group (i.e. products to which the DRCM applies at time 0) and the control group (the remaining products). The solid and dashed vertical lines indicate the date at which the DRCM starts to apply and the date at which the reform was announced.

FIGURE 3.B.9: MEAN GAPS PER GROUP OVER TIME — EVENT IN IMPORTING COUNTRY (CONTINUES IN NEXT FIGURE)



No: 37, Side: Importer, Country: LV, Date: 4.2016

No: 38, Side: Importer, Country: LV, Date: 7.2016

No: 39, Side: Importer, Country: LV, Date: 1.2017

No: 40, Side: Importer, Country: LV, Date: 1.2018

No: 41, Side: Importer, Country: NL, Date: 1.2007

No: 42, Side: Importer, Country: NL, Date: 4.2013

No: 43, Side: Importer, Country: PL, Date: 7.2011

No: 44, Side: Importer, Country: PL, Date: 10.2013

No: 45, Side: Importer, Country: PL, Date: 7.2015

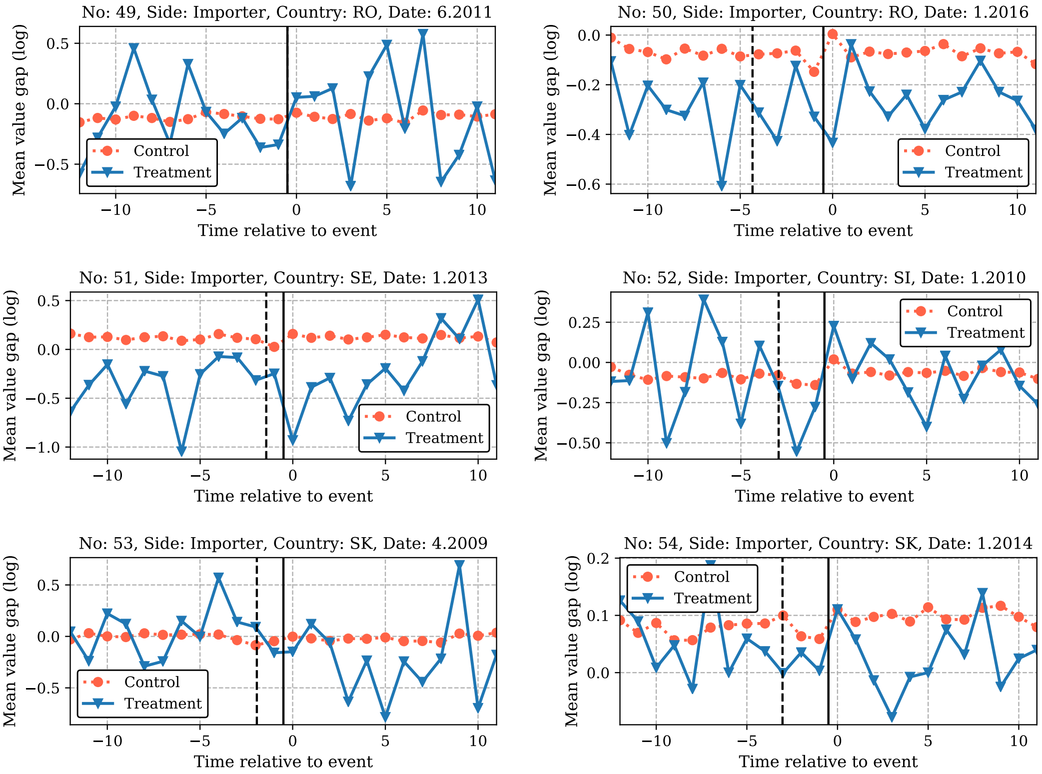
No: 46, Side: Importer, Country: PL, Date: 1.2017

No: 47, Side: Importer, Country: PL, Date: 1.2018

No: 48, Side: Importer, Country: PT, Date: 10.2006

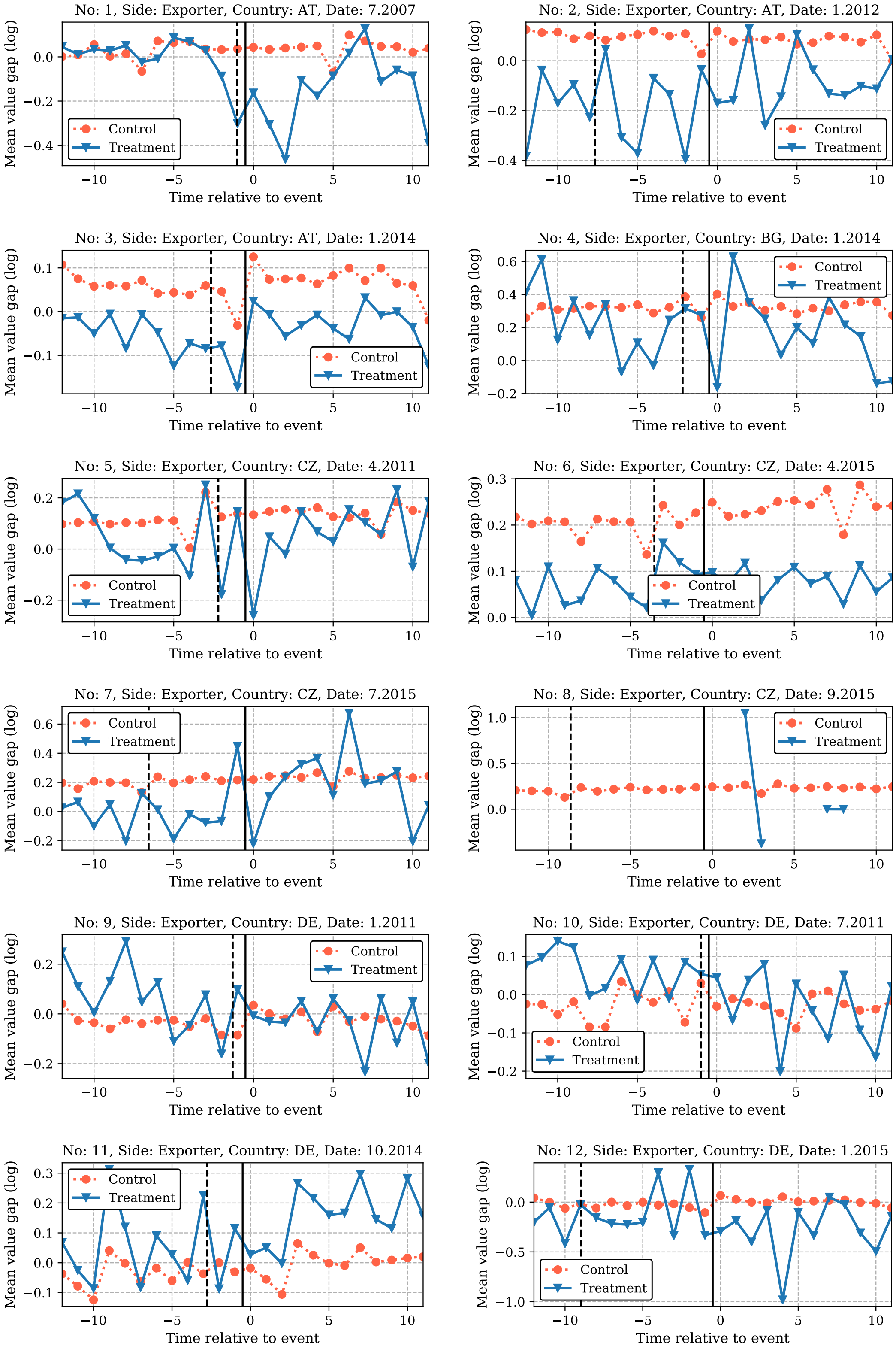
Note: Time series of the mean of gaps as calculated in (3.1), for the treatment group (i.e. products to which the DRCM applies at time 0) and the control group (the remaining products). The solid and dashed vertical lines indicate the date at which the DRCM starts to apply and the date at which the reform was announced.

Figure 3.B.10: Mean gaps per group over time — Event in importing country (continued from previous figure)



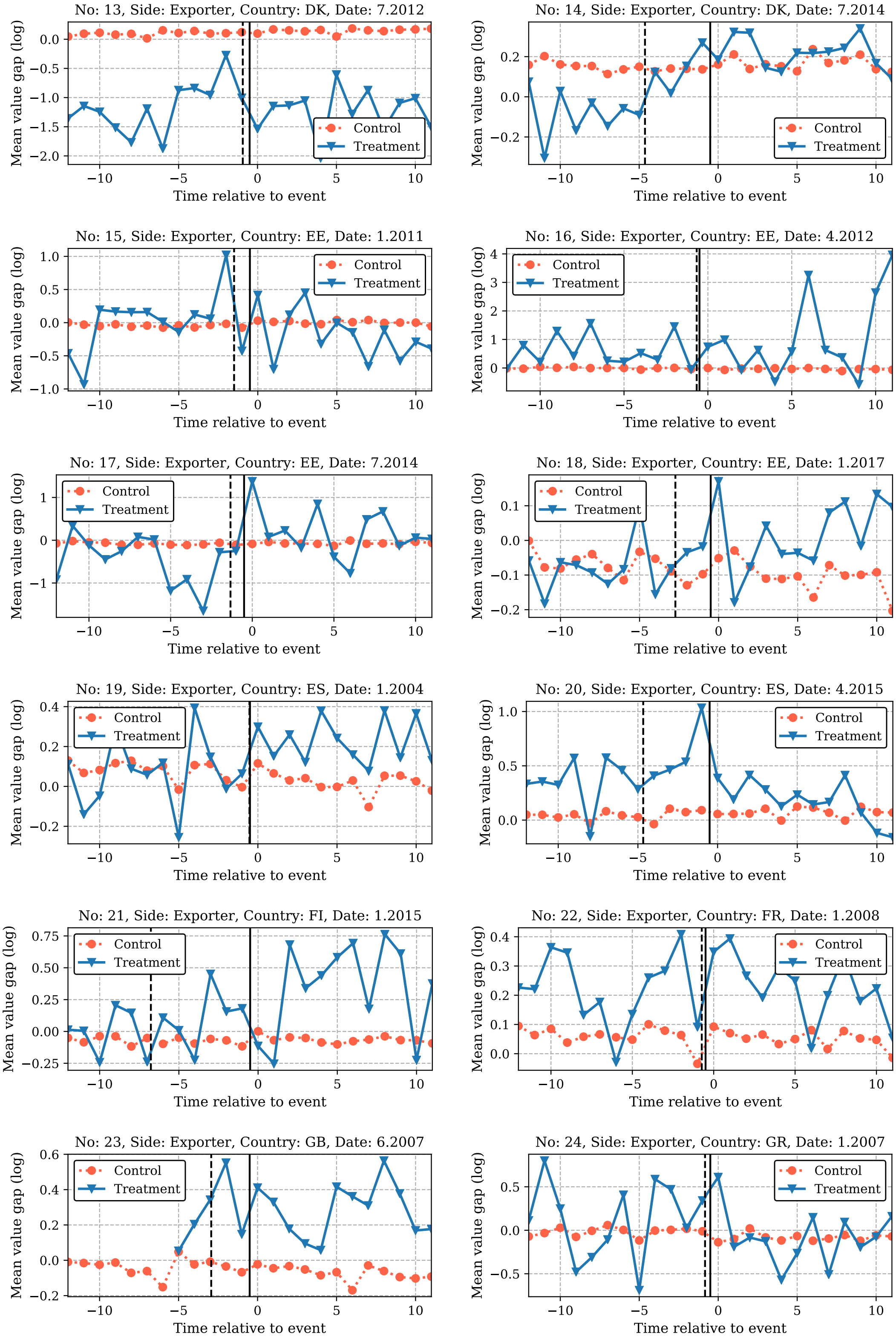
Note: Time series of the mean of gaps as calculated in (3.1), for the treatment group (i.e. products to which the DRCM applies at time 0) and the control group (the remaining products). The solid and dashed vertical lines indicate the date at which the DRCM starts to apply and the date at which the reform was announced.

FIGURE 3.B.11: MEAN GAPS PER GROUP OVER TIME — EVENT IN EXPORTING COUNTRY (CONTINUES IN NEXT FIGURE)



Note: Time series of the mean of gaps as calculated in (3.1), for the treatment group (i.e. products to which the DRCM applies at time 0) and the control group (the remaining products). The solid and dashed vertical lines indicate the date at which the DRCM starts to apply and the date at which the reform was announced.

Figure 3.B.12: Mean gaps per group over time — Event in exporting country (continues in next figure)


No: 13, Side: Exporter, Country: DK, Date: 7.2012

No: 14, Side: Exporter, Country: DK, Date: 7.2014

No: 15, Side: Exporter, Country: EE, Date: 1.2011

No: 16, Side: Exporter, Country: EE, Date: 4.2012

No: 17, Side: Exporter, Country: EE, Date: 7.2014

No: 18, Side: Exporter, Country: EE, Date: 1.2017

No: 19, Side: Exporter, Country: ES, Date: 1.2004

No: 20, Side: Exporter, Country: ES, Date: 4.2015

No: 21, Side: Exporter, Country: FI, Date: 1.2015

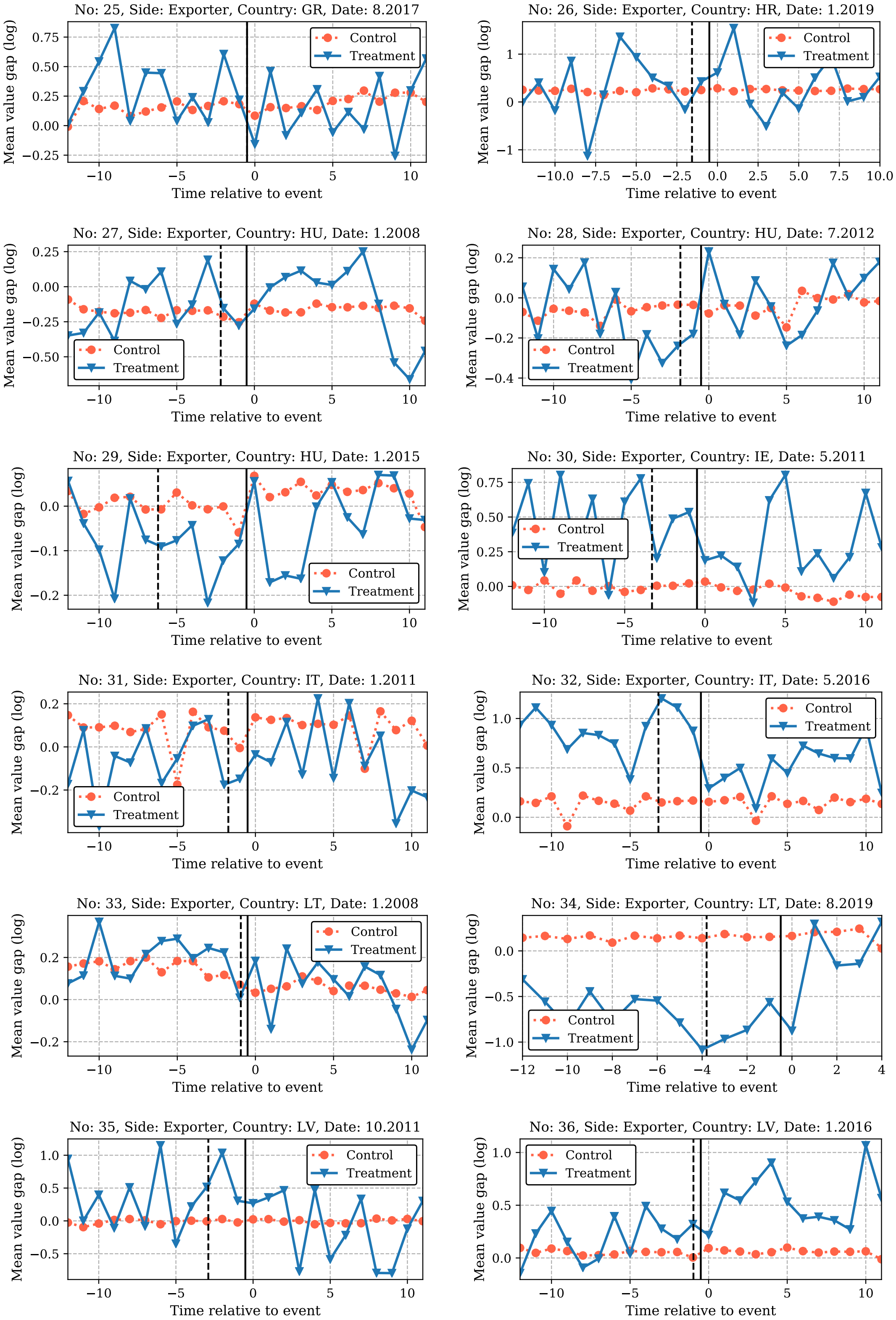
No: 22, Side: Exporter, Country: FR, Date: 1.2008

No: 23, Side: Exporter, Country: GB, Date: 6.2007

No: 24, Side: Exporter, Country: GR, Date: 1.2007

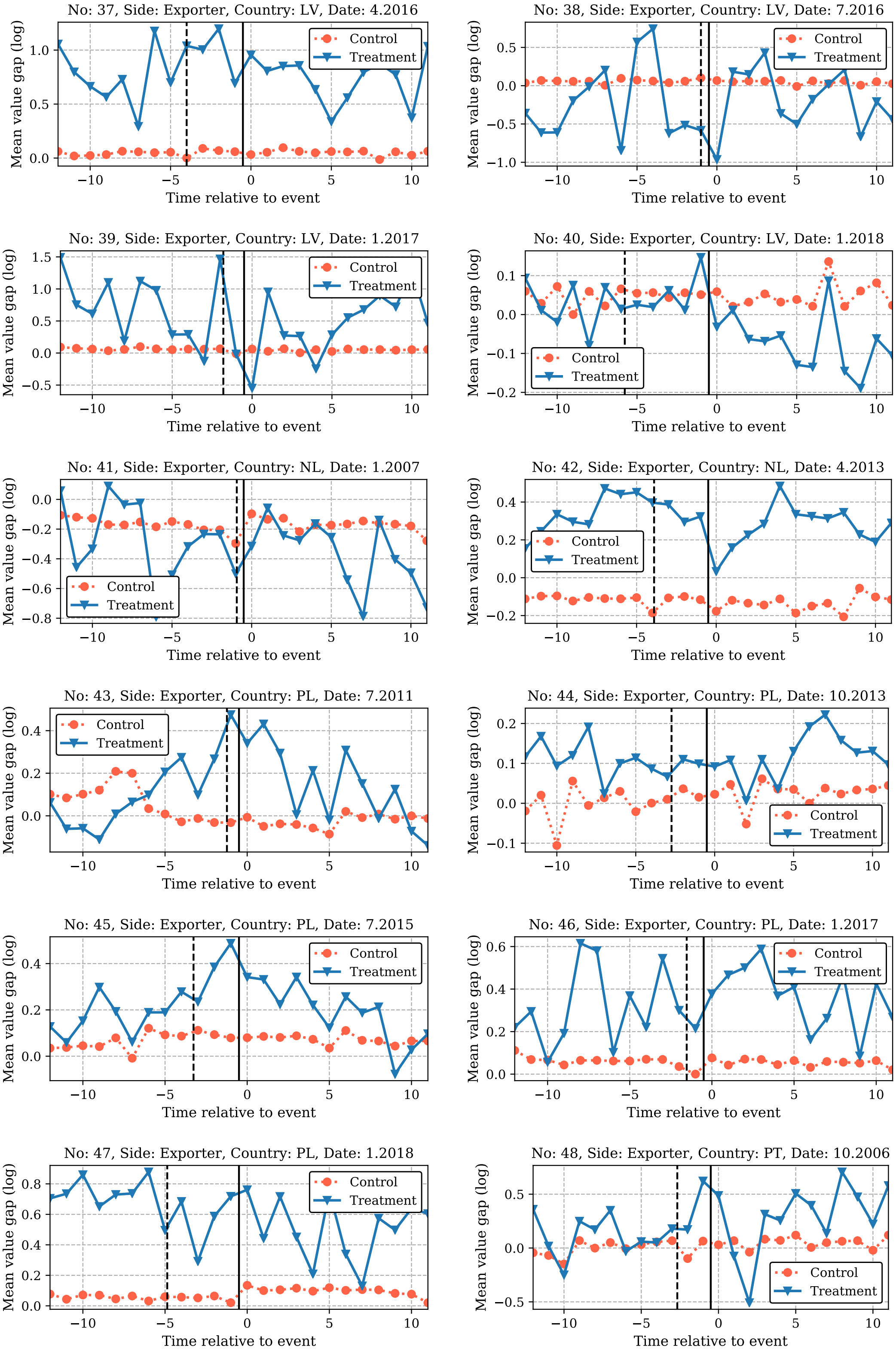
Note: Time series of the mean of gaps as calculated in (3.1), for the treatment group (i.e. products to which the DRCM applies at time 0) and the control group (the remaining products). The solid and dashed vertical lines indicate the date at which the DRCM starts to apply and the date at which the reform was announced.

FIGURE 3.B.13: MEAN GAPS PER GROUP OVER TIME — EVENT IN EXPORTING COUNTRY (CONTINUES IN NEXT FIGURE)



Note: Time series of the mean of gaps as calculated in (3.1), for the treatment group (i.e. products to which the DRCM applies at time 0) and the control group (the remaining products). The solid and dashed vertical lines indicate the date at which the DRCM starts to apply and the date at which the reform was announced.

Figure 3.B.14: Mean gaps per group over time — Event in exporting country (continues in next figure)


No: 37, Side: Exporter, Country: LV, Date: 4.2016

No: 38, Side: Exporter, Country: LV, Date: 7.2016

No: 39, Side: Exporter, Country: LV, Date: 1.2017

No: 40, Side: Exporter, Country: LV, Date: 1.2018

No: 41, Side: Exporter, Country: NL, Date: 1.2007

No: 42, Side: Exporter, Country: NL, Date: 4.2013

No: 43, Side: Exporter, Country: PL, Date: 7.2011

No: 44, Side: Exporter, Country: PL, Date: 10.2013

No: 45, Side: Exporter, Country: PL, Date: 7.2015

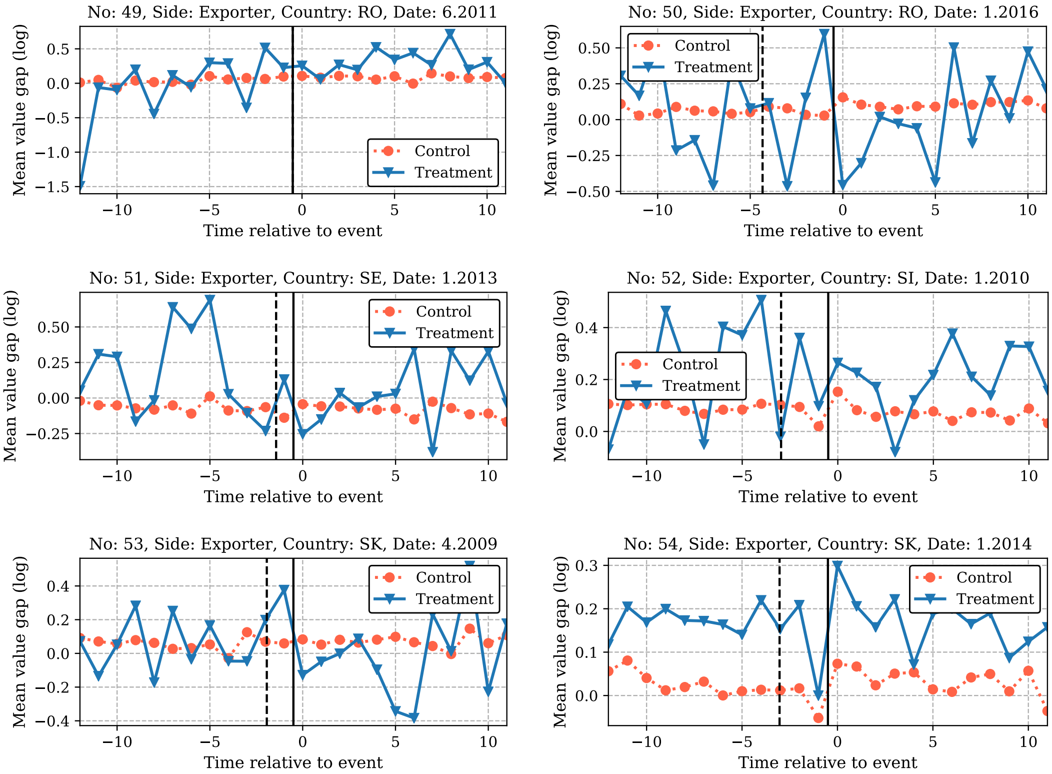
No: 46, Side: Exporter, Country: PL, Date: 1.2017

No: 47, Side: Exporter, Country: PL, Date: 1.2018

No: 48, Side: Exporter, Country: PT, Date: 10.2006

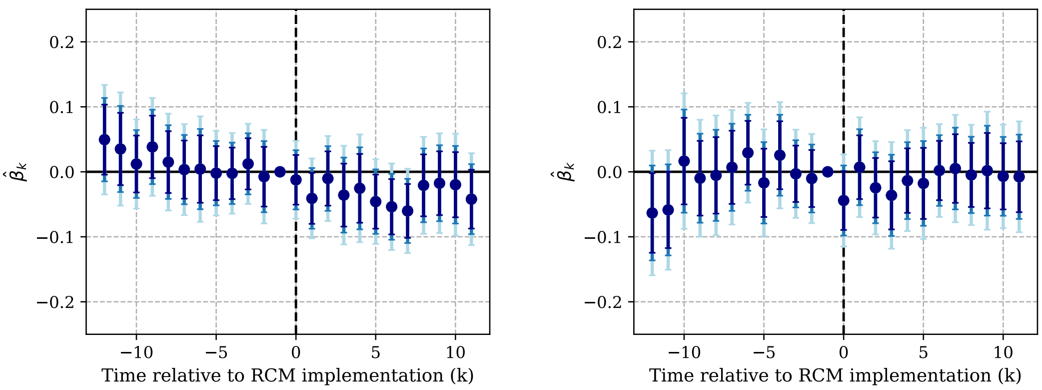
Note: Time series of the mean of gaps as calculated in (3.1), for the treatment group (i.e. products to which the DRCM applies at time 0) and the control group (the remaining products). The solid and dashed vertical lines indicate the date at which the DRCM starts to apply and the date at which the reform was announced.

FIGURE 3.B.15: MEAN GAPS PER GROUP OVER TIME — EVENT IN EXPORTING COUNTRY (CONTINUED FROM PREVIOUS FIGURE)



Note: Time series of the mean of gaps as calculated in (3.1), for the treatment group (i.e. products to which the DRCM applies at time 0) and the control group (the remaining products). The solid and dashed vertical lines indicate the date at which the DRCM starts to apply and the date at which the reform was announced.

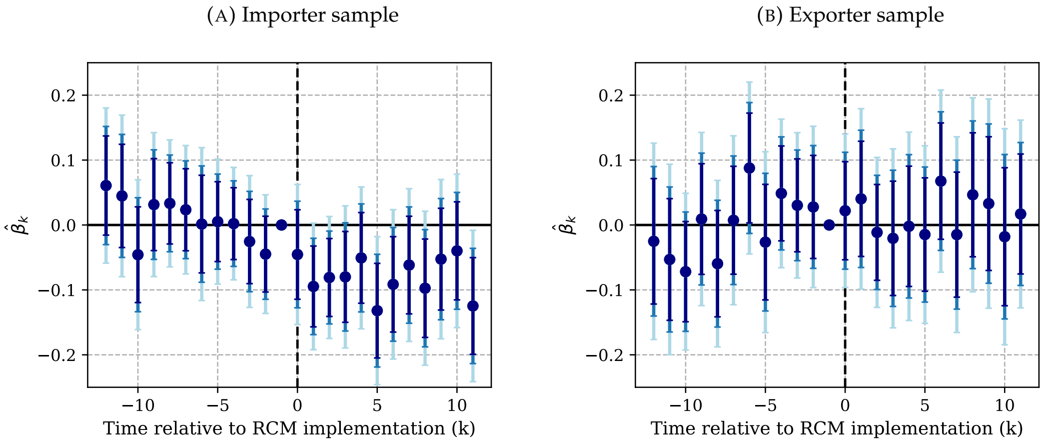
FIGURE 3.B.16: DYNAMIC DIFFERENCE-IN-DIFFERENCE: QUANTITY GAP


(A) Importer sample

(B) Exporter sample

Note: Sequences of $\hat{\beta}_k$ from equation (3.3), estimated on the importer and exporter samples with the following fixed effects: $j, t, ymc, ymkp$. The coefficient in the period immediately preceding the implementation of the DRCM has been normalized to 0. 99%-, 95%- and 90%-confidence intervals are depicted in different colour shades.

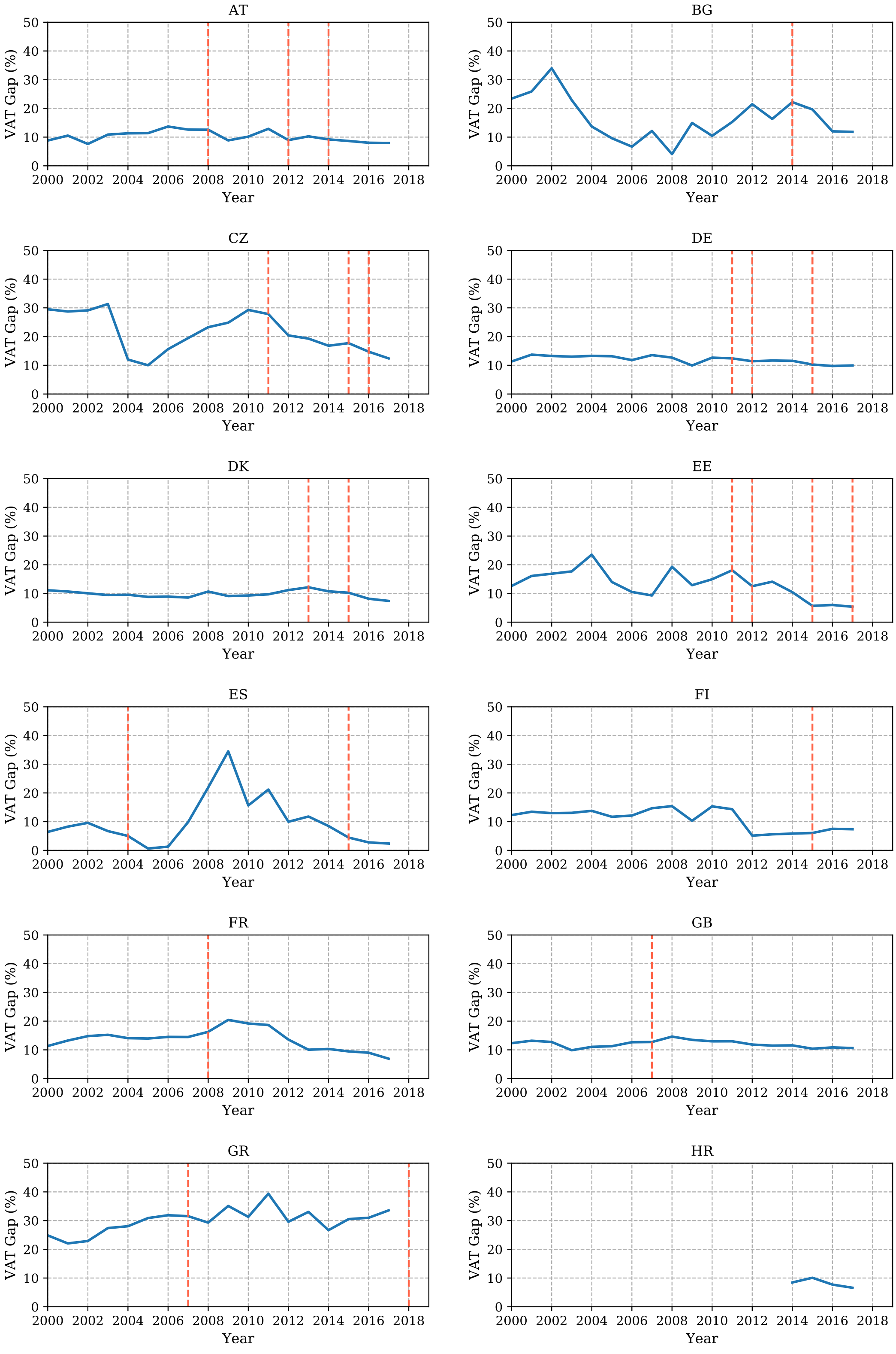
FIGURE 3.B.17: DYNAMIC DIFFERENCE-IN-DIFFERENCE: VALUE GAP — RESTRICTED CONTROL GROUP


(A) Importer sample

(B) Exporter sample

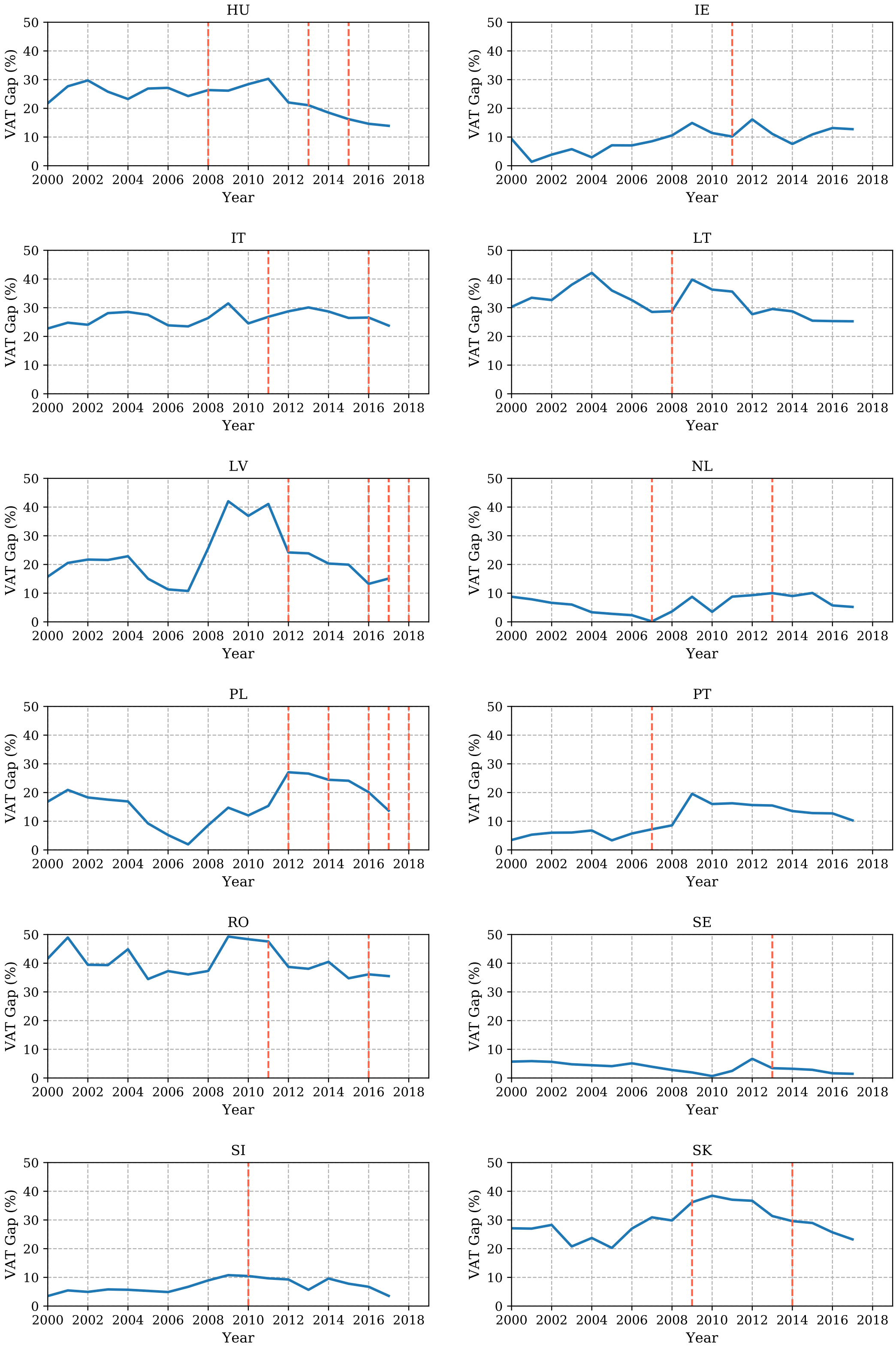
Note: Sequences of $\hat{\beta}_k$ from equation (3.3), estimated on the importer and exporter samples with the following fixed effects: $j, t, ymc, ymkp$. The coefficient in the period immediately preceding the implementation of the DRCM has been normalized to 0. 99%-, 95%- and 90%-confidence intervals are depicted in different colour shades.

FIGURE 3.B.18: VAT GAPS OVER TIME IN EACH MEMBER STATE (CONTINUES IN NEXT FIGURE)



Note: VAT gaps expressed in percentage of the theoretical tax liability. The vertical lines denote the times of implementation of the DRCM.

Figure 3.B.19: VAT gaps over time in each Member State (continued from previous figure)



Note: VAT gaps expressed in percentage of the theoretical tax liability. The vertical lines denote the times of implementation of the DRCM.

Appendix 3.C Instances of the application of the DRCM to goods in EU Member States

The reverse charge mechanism can be applied domestically by EU Member States to certain categories of goods that are specified in articles 199 and 199a-c of Council Directive 2006/112/EC. There may also be exceptions not based of the aforementioned articles whereby the DRCM is applied by member states, a list of which can be found in table 1 of de la Feria (2019). The data collection procedure I followed consists in checking notifications by Member States informing the European Commission of their use of the DRCM (an up-to-date list entitled *Notifications of the VAT committee* can be found online, last accessed 19.05.2020). Some events were found using other sources, e.g. European Commission (2014) or de la Feria (2019). Once the list of events was established, I systematically consulted the national laws to find details of the timing (implementation and announcement dates) and the coverage (i.e. the CN codes to which the DRCM applies).

Austria The DRCM has been applied to mobile phones and electronic integrated circuits since 01.01.2012. Although some sources suggest it may have been implemented earlier (de la Feria, 2019), documents from news and the official European Commission documentation suggests the above date is the actual one.[36] The DRCM has been extended to game consoles, laptop computers and selected supplies of raw and semi-finished metals, of which a list can be found in European Commission (2014, p. 92). The DRCM is also applied to waste and scraps since 01.07.2007 (list of goods available in BGBl. II No. 129/2007).

Bulgaria A selection of 41 cereals are subject to the DRCM. It was implemented from 01.12.2013, and the full list of CN codes subject to the DRCM can be found in Part Two of Annex 2 to Chapter 19(a) (art 163(a)) of the VAT Act (602-01-23 / 12.04.2006).[37] The DRCM also applies to categories of waste as defined in Part One of Annex 2 to Chapter 19(a) of the VAT Act, effective 01.01.2007 (State Gazette, issue 108 of 29.XII).

Croatia In Croatia, the DRCM applied to recyclable waste part processed waste and other processed industrial materials. However, it applied from the date of its accession to the EU. Therefore, there is no pre-treatment data available. The DRCM also applies to certain steel products for concrete reinforcement from 01.01.2019 (art. 10 of OG 106/2018 modifying art. 75 of the Value Added Tax Act).

Czech Republic The first implementation of the DRCM dates from 01.04.2011, as specified by art 199 of the EU VAT Directive. It applied to scraps and wastes — metal, paper, etc. (Grásgruber et al., 2013). The full list of these products can be found in Annex 5 of Act No. 235/2004 Coll article 92c, the law on VAT in Czech Republic.

---

[36] See the International VAT Monitor July/August 2011, and the annex to COM/2018/0118 final (last accessed 19.05.2020).

[37] Which can be accessed here, last accessed 21.02.2020.

Starting 01.04.2015, the DRCM has been applied to mobile phones, electronic integrated circuits, laptop computers and game consoles, and a variety of crops and supplies of metals. Its scope was extended on two occasions in the following months: to selected agricultural crops (01.07.2015) and to sugar beets (01.09.2015). A General Reverse Charge Mechanism (GRCM) is being implemented following EU Council Directive 2018/2057, starting January 2020.

Denmark The DRCM was applied to domestic transactions of mobile phones, electronic integrated circuits, laptop computers and video consoles from 01.07.2014. It only applies when the provider's activity is not primarily selling these goods. The firm who has the burden of the tax is specified in art 46(1), nos 8-10 of the VAT law. The DRCM also applies to the supplies of scrap metal from 01.07.2012 (art. 3 of Law no. 590 of 18/06/2012).

Estonia The first instance of application of the DRCM to goods in Estonia dates from 01.01.2011 when it was applied to metal waste (RT I, 10/12/2010, 3). It was then applied to gold effective 01.04.2012 (RT I, 27.03.2012, 7, amending art. 41(1) of the VAT Act RT I 2003, 82, 554). The relevant list of metal waste can be found in art. 104 of the Waste Act (RT I 2004, 9, 52). It was further extended to precious metals effective 01.07.2014 (RT I, 06.06.2014, 2). Last, the DRCM was then applied to selected metals on 01.01.2017 (RT I, 08.11.2016, 1) and this selection was later very slightly modified (RT I, 24.04.2018, 2).

Finland The DRCM applies only to goods from metal waste, effective 01.01.2015 (27/06/2014/507 modifying art 8d of the 30.12.1993/1501 VAT Act).

France The goods subject to the DRCM are given in art. 283 CGI. Since 01/01/2008, it applies to new industrial waste and recoverable materials (LOI n 2007-1824 du 25 décembre 2007 - art. 57).

Germany The DRCM was first introduced on selected used materials (metals, described in Annex 3 of the VAT law) and gold on 01.01.2011 (Jahressteuergesetz 2010, amending article 13b of the VAT law). It was followed by mobile phones and electronic integrated circuits on 01.07.2014 (Artikel 6 G. v. 16.06.2011 BGBl. I S. 1090), video consoles and tablets (excluding computers) on 01.10.2014 (Artikel 8 G. v. 25.07.2014 BGBl. I S. 1266). On that occasion, a list of metals was also included, presented in Annex 4 of the law. This list was further amended effective 01.01.2015 (Artikel 11 G. v. 22.12.2014 BGBl. I S. 2417).

Greece The details of which goods are subject to the DRCM are contained in article 39a of Law No. 2859/2000. The DRCM was introduced to mobile phones, video consoles and laptop computers effective 01.08.2017 (article 67 of Law 4484/2017). It also applied to supplies of recyclable waste since 01.01.2007 (paragraph 2 of article 21 of law 3522/2006).

Hungary The details are contained in art. 142 of the Act CXXVII of 2007 on Value Added Tax, and the annexes referred in art. 142 (d), (i) and (j). Chronologically, the DRCM entered into force as follows. For the sale of waste: 01.01.2008 (original law); for the sale of selected agricultural commodities: 01/07/2012 (Act XLIX of 2012); for the sale of selected metals: 01/01/2015 (Act XXXIII of 2014).

Ireland The DRCM applied to scrap metal effective 01.05.2011 (art 16 of Value-Added Tax Consolidation Act 2010, modified by the Finance Act 2011 (No. 6 of 2011), ss. 59(1)(a), 59(2)).

Italy The VAT reverse-charge provisions are described in art. 17 of Presidential Decree No. 633 of 1972. It applied to computers and video consoles from 02.05.2016 (Legislative Decree No. 24 of 11 February 2016). Before that, it has applied to mobile phones and electronic integrated circuits effective 01.01.2011, as authorized by the EU Council (2010/710/EU). The DRCM also applies to waste (art. 74) since at least 2003, but these supplies appear not to have been taxed before. Due to the uncertainty relative to the tax treatment of these supplies, they are excluded from the sample.

Latvia The articles covered by the DRCM are contained in articles 141 to 143 of the VAT law (2012/197.2). It was applied from 01.04.2016 to mobile phones, electronic integrated circuits and laptops (2015/248.18), from 01.07.2016 to a selection of cereals (2016/120.2) and from 01/01/2017 to precious metals (2016/241.48). Effective from 01.01.2018 to video game consoles, metal products and the supply of household electronic equipment and electrical household appliances (2017/156.11, see annex to the law for lists of products in 2013/14.3, see versions for relevant dates). The list of metal products was shortened effective 01.07.2019 and the DRCM on the supply of household electronic equipment and electrical household appliances was abolished effective 01.01.2020 (2019/133.4). The DRCM also applies to wood timber, effective 01.07.1999 (Latvijas Vēstnesis, 133/135, 30.04.1999; adding article 13(2) to the previous VAT law: Latvijas Vēstnesis , 49, 30.03.1995), and to other types of forms of wood products, effective 01.01.2016 (2015/248.18). Last, it applies to metal waste and scraps, effective 01.10.2011 (Latvijas Vēstnesis, 117, 28.07.2011.).

Lithuania The details on which products are subject to the DRCM is not directly contained in the VAT Act, but in the *No. 900 Amendment To The Measures To Ensure Tax Liability*. The DRCM applies to scrap metals as well as selected timber products effective 01.01.2008 (Resolution No. 1390 of the Government of the Republic of Lithuania of 19 December 2007). It became mandatory for mobile phones, tablets, and portable laptop computers as well as hard disks from 01.08.2019 (Resolution No. 6962 of the Government of the Republic of Lithuania of April 24, 2019).

The Netherlands Details on the application of the DRCM are specified in the Implementation Decree on turnover tax 1968 (Uitv. Besl. OB 1968, articles 24ba and 24bb). The DRCM was first applied to used materials, scrap and waste from 01.01.2007

(Stb. 2006, 684) and then applied to mobile phones, chips, game consoles, tablet computers and laptops from 01.04.2013 (stb-2012-694). It is worthwhile noting that for these goods (except computers) the DRCM has been optional since 01.06.2012, and became mandatory as of 01.04.2013. The DRCM has been applied to clothing other than footwear over the period 1992-2012 (de la Feria (2019) and officially announced in OJ L 351, 2 December 1992, renewed in OJ L 8, 14.1.1998, p. 16 and 2007/740/EC). The DRCM on clothing ceased to operate on 01.04.2013 (stb-2012-694).

Poland The DRCM is abolished in Poland from 01.11.2019, date at which it is replaced by a "split payment" mechanism. Before that, goods subject to reverse charge were specified in Annex 11 (but the goods are classified using the Polish classification, and must be translated into the CN nomenclature manually), as referred in art. 17.1.8 of the VAT Act (Dz. U. z 2004 r. nr 54, poz. 535). This Annex was modified over time as follows. Effective 01.07.2011, it contained a list of metal scraps (Dz. U. z 2011 r. nr 134, poz. 780); it was then extended to more such products on 01.10.2013 (Dz. U. z 2013 r. poz. 1027); then to more metals as well as laptops, game consoles and mobile phones on 01.07.2015 (Dz. U. z 2015 r. poz. 605) and finally to computer chips and other metals from 01.01.2017 (Dz. U. z 2016 r. poz. 2024). The DRCM also applies to hard disks since 01.01.2018 by derogation allowed by the EC (2017/0205 (NLE)).

Portugal In portugal, Annex E of the VAT Code, referred to in article 2 1(i) of the Value Added Tax Code, contains the list of supplies (waste and metals) to which the DRCM applies effective 01.10.2006 (Lei n. 33/2006). The DRCM also applied to doorsteps as a derogation between 2004 and 2015 (de la Feria, 2019), but I could not find evidence of it in the laws.

Romania The VAT rules are contained in title VII of law no. 227/2015, and the goods for which DRCM applies are detailed in art. 331 (and in art. 160 of law 571/2003 before 2015). The DRCM was applied to a selection of cereals effective 31.05.2011, as specified in EMERGENCY ORDER no. 49 of May 31, 2011. It also applies to mobile phones, electronic integrated circuits, video console, laptops and tablets effective 01.01.2016 (LAW no.227 of September 8, 2015 on the Fiscal Code). The DRCM also applies to metal scraps and other waste, as well as to wood. However, the application of the DRCM pre-dates the joining of Romania to the EU, and I could not find the CN codes matching these goods (art 331 a and b).

Slovakia The goods to which the DRCM apply are listed in art. 69(12) of the Act No. 222_2004 Coll. This article has been amended over time to include mobile phones and electronic integrated circuits, selected cereals and articles of iron and steel (360/2013 Coll., effective 01/01/2014). Before that, the DRCM has been applied to metal waste and scraps as well as gold (83/2009 Coll., effective 01.04.2009).

Slovenia The goods to which DRCM apply are specified in article 76a of the Value Added Tax Act, and refer to Annex III for a list of goods. The DRCM applies to waste and scrap since 01.01.2010 (The Act Amending the Value Added Tax Act - ZDDV-1B

(Official Gazette of the Republic of Slovenia, No. 85/2009, of 30 October 2009)). The list was modified on 01.01.2011, restricting the goods to which the DRCM applies.

Spain VAT regulations are specified in Law 37/1992, of December 28, on Value Added Tax. Article 84 dictates which goods are subject to reverse charge. The DRCM applies to supplies of mobile phones, video game consoles, laptops and digital tablets, as well as selected metals since 01.04.2015 (Art. 1.19 of Law 28/2014, a full list of the goods is available in section 10 of the Annex of Law 37/1992). It also applies to different types of industrial waste and metals since 01.01.2004 (Art. 7.3.2 of Law 62/2003, full list in Annex 7).

Sweden The DRCM applies to metal scrap and waste, as specified in art. 2(6). The list of CN codes is directly in the law, and it is effective 01.01.2013 (SFS 2012:755 Law on amendments to the VAT).

United Kingdom Only mobile phones and electronic integrated circuits are subject to the DRCM in the UK, effective 01.06.2007 (2007 No. 1417).

Other EU countries In Belgium , Malta and Luxembourg, the DRCM only applies to some services and intangibles (e.g. CO_2 emission permits, construction, electricity, etc.).

Instances of the DRCM not included in the analysis There are several instances where the DRCM was applied, but the episode is dropped from the sample. Cyprus: it applies to waste and scraps, but only for firms from a specific industry. Portugal: the DRCM applied to doorsteps, but no evidence of it could be found in the law. Romania: the DRCM applied to waste and timber since the time at which the country joined the EU, eliminating the possibility to observe the treated products prior to the application of the DRCM. Italy: the DRCM applies to waste, but it appears that these goods were not subject to VAT prior to the DRCM implementation. Croatia: the DRCM applied to waste and gold since the time at which the country joined the EU, eliminating the possibility to observe the treated products prior to the application of the DRCM. Idem for Bulgaria (01.01.2007), and the Netherlands (01.12.1992) and Latvia (1999, not in the EU even in the post period).