Stockholm School of Economics
Department of Economics
5350 Master's Thesis in Economics
Academic year 2022–2023

Mind the Gap: Quantifying fraud on Swedish exports using bilateral trade asymmetries

Giulian Etingin-Frati (42146) & Ivelina Ilkova (42134)

Abstract: Using open-source international trade statistics, we analyse cases where export country statistics do not match import country statistics for the same trade flow due to fraud. Specifically, we are interested in uncovering undervaluation fraud, whereby importers understate the value of their goods in order to pay lower taxes. Using a pooled cross-section approach, we confirm that previous findings about the salience of the tariff rate apply to Swedish exports. We then introduce covariates for government type and quality in countries which import Swedish goods. We find that both low democracy and high corruption contribute to higher incidence of undervaluation fraud, and that the duty rate effect is more pronounced in less democratic countries. To our knowledge, we are the first researchers to investigate this issue from the Swedish export perspective. We discuss the possible mechanisms behind these effects and present their full investigation as an avenue for further research.

Keywords: International taxation, trade, tax evasion, customs fraud
JEL: F14, H26

Supervisor: Céline Zipfel
Date submitted: 14 May 2023
Date examined: 26 May 2023
Discussants: Cajsa Klaß and Milan Spendel
Examiner: Karl Wärneryd

We would like to acknowledge the valuable contributions and kind advice of Thomas Fije McAdam and Adrien Bussy.

Contents

1 Introduction

Taxes are a key way of raising revenue for governments. In OECD countries, taxes amounted to 34.11% of GDP on average in 2021, the latest year for which statistics are available (OECD, 2023). While they contribute to government coffers, taxes also present revenue loss risks for governments, since dishonest taxpayers will attempt to evade them, reducing government revenues. Quantifying losses from evasion is difficult due to their criminal nature, but we know that these can be significant. In 2021, global losses due to tax evasion by corporations and individuals were estimated at approximately $483B USD (Tax Justice Network, 2021). This lost revenue remains in private hands or is directed to the personal accounts of government officials, bypassing the state and resulting in lower government spending. Countries with lower state capacity, or which do not effectively enforce the rule of law, are particularly at risk of higher rates of tax evasion, since evaders face relatively low consequences. Given that this is more likely to occur in developing countries which could materially benefit from increased revenues, there is a clear interest in not only quantifying current and historic levels of tax evasion, but also understanding the circumstances in which it occurs. If there is indeed a link between country characteristics, such as their levels of democracy, law enforcement, and endemic corruption, and tax evasion, then these could be targeted for improvements. The successful implementation of such a program would increase government revenues, reduce reliance on foreign aid, and increase self-sufficiency in developing countries.

Due to the myriad taxes levied by countries around the world, each with its own tax rates, tax bases, and implementation rules, we narrow our analysis to import taxes. This is because in low- and middle- income countries, indirect taxes (i.e. taxes which are levied on a per transaction basis) make up a higher proportion of tax revenues than in high-income countries since they are easier to collect than direct taxes (i.e. those levied on income, which require more information and costlier enforcement (Martinez-Vazquez et al., 2011)). In high-income countries, taxes on international trade made up approximately 0% of tax revenue in 2020. In low- and middle-income countries, the figure was 4.8%, and in Sub-Saharan Africa it was 9.1% (World Bank, 2023). Depending on the country of import, a number of taxes can be levied on imported goods. The most common of these are customs duties, excise duties, and Value Added Tax (VAT). Since our main goal is to understand the drivers of revenue losses due to tax evasion in developing countries, customs duties are an important place to start.

Following the majority of the literature, we choose to focus on customs duties exclusively in our analysis below. This is for a number of reasons. Firstly, customs duties are levied on a wide range of goods and can in general not be reclaimed by the importer. They therefore represent a fixed cost charged by customs authorities. Conversely, VAT is also levied at import but can usually be reclaimed by the importer if the importer is not the final consumer of the goods (i.e. a firm importing goods for resale). In these cases, import VAT represents a cashflow issue rather than a cost and is therefore less likely to strongly influence importer behaviour. While

excise duties are similar to customs duties in that they are levied on import and cannot be reclaimed, they are mostly applied to a small number of specific goods such as alcohol, tobacco, and vehicles, with the tax base varying from country to country. Since VAT is most often reclaimed and excise duties are levied on very specific products, they are less likely to accurately capture the effect of tax changes on global customs fraud. Therefore, our analysis will focus on customs duties alone, as they are the most salient tax out of the three.

To conduct our analysis, we use a technique called "mirror analysis", which has been in use since the 1960s to quantify international trade fraud. To our knowledge, we are the first researchers to use the technique on Swedish exports. In essence, this involves comparing statistics published by import and export countries for the same trade flows. In theory, countries on the import and export sides of the same flow should report the same data for the weight and values of the goods traded between them. While some disparities in the statistics are to be expected (as explained in more detail in Section 2.3.1), systematic and persistent "gaps" may be evidence of fraud, especially on the import country side. For simplicity, we mainly focus on Swedish exports to all countries for the period from 2000 - 2021. Using the UN COMTRADE trade statistics database, we are able to gather data on both sides of these trade flows, along with other data on trade and country characteristics from the World Trade Organisation, the World Bank, and the VDEM dataset.

Our results suggest that duty rates have a highly significant effect on trade gaps, indicating that customs fraud is likely to be present, and that it is especially likely to occur in trade flows of highly dutiable products. Further, we find that this relationship is more pronounced in import countries that have worse democracy and corruption scores. These effects are robust to the inclusion of a number of controls, as well as import country and year fixed effects. We further present findings that these effects hold not only for Swedish exports, but also for exports from Denmark, Norway, and Finland. We therefore confirm that the findings of the previous literature on this topic apply to the Nordic countries, and propose further research into the mechanisms behind our results.

The structure of our thesis is as follows: Section 2 explains the import and export processes, as well as common import tax evasion methods, Section 3 discusses the literature and our contribution to it, Section 4 outlines the data used, Sections 5 and 6 contain the empirical strategy and discussion of results, and Section 7 concludes.

2 Background

2.1 Trade practicalities

Since we focus on exports from Sweden, we will outline the administrative aspects of the trade process from this perspective. The export process begins when Swedish goods are ordered by foreign customers. This procedure is presented in a stylised fashion in Figure 1. We note here that the processes for export to EU countries and non-EU countries are slightly different, and therefore we will explain each separately.

When sending goods outside of the EU, Swedish exporters must first register their exports with the Swedish customs authority, *Tullverket*. This is done through the filing of an "export declaration", which contains details such as the identity of the exporter, the type of goods, their quantity, value, and destination country. Tullverket processes these documents electronically, selecting a small number of shipments to physically inspect using a risk-based approach (Hörnqvist, 2006). The vast majority of shipments however are not physically inspected (European Commission, 2023). Once the documents are accepted by the customs authorities, the goods can leave Sweden and the export process is complete (European Commission, 2021).

On arrival of the goods in the import country, the importer files an "import declaration" with the customs authorities in that country, usually containing very similar information as the export declaration filed in Sweden earlier in the process. The goods can be inspected by the import country customs authorities at this stage. Theoretically, import country customs authorities should ensure that the details given at export and at import match, though this may not always happen. Once the import declaration is accepted by the authorities, it is also used to determine any import taxes due. Having paid any applicable taxes, importers are free to take possession of their goods.

For shipments within the EU, the process is much simpler since, from a customs perspective, there is no difference between shipments within Sweden and those to other EU member countries (European Commission, 2020). These shipments do not attract duties, and also are not subject to the customs formalities explained above. Nonetheless, importers and exporters must still report their trade flows for statistical purposes. While there could be some implications for controlled (e.g. dangerous or military) goods or VAT collection, for the purposes of our analysis, we assume that this trade is frictionless and that there are no duties levied on any trade with EU countries from Sweden.

2.2 Harmonised System (HS)

In WTO member countries, the Harmonised System (HS) is used to classify traded goods (Yu, 2008). It is a hierarchical taxonomy of numerical codes which are used to specify the precise nature of the goods being traded. Conveniently for our analysis, these codes are standard across the World Trade Organisation (WTO) and are also the basis on which customs duties are levied.

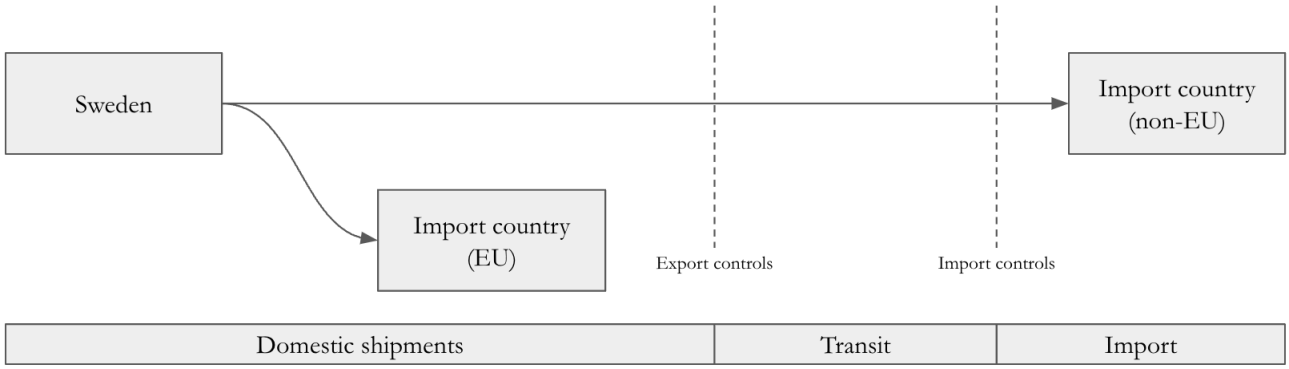

Figure 1: Stylised trade process

The HS is divided into 99 "chapters", which represent broad categories of goods. For example, chapter 01 is used for live animals, chapter 16 for preparations of meat, fish, and seafood, and chapter 57 for carpets and other floor coverings. Since the chapter numbers have two digits, this is termed the "HS2" level of granularity. Each chapter is further subdivided into higher levels of granularity. For example, within chapter 01, heading 0102 is used for live bovine animals, and 010221 is used for purebred live bovine animals imported for breeding purposes. The four- and six-digit headings comprise the "HS4" and "HS6" levels of granularity, respectively.

In total, there are 96 subdivisions at the HS2 level, 1,224 at the HS4 level, and 6,533 at the HS6 level. While some codes are retired and others added over time to account for changes in products on the global market, the HS is largely stable, enabling comparisons across both time and product types. We specifically address concerns about the consistency of the HS in Section 2.3.3. Throughout our analysis, we mostly rely on the HS4 level of granularity, for reasons we explain in Sections 2.3.3 and 5.3.1.

2.3 Explanation for gaps

2.3.1 Benign reasons

Import statistics are usually reported on the Cost, Insurance, and Freight (CIF) basis, meaning that they include the value of the goods along with the costs of transporting them from the export country and any insurance paid for this transport. Conversely, export statistics are usually calculated on the Free on Board (FOB) basis, meaning that they include only the value of the goods (Bussy, 2020; Markhonko, 2014). This mismatch could be the cause of some trade gaps seen in our analysis. We note that a very small number of countries deviate from these conventions, using the FOB basis for both imports and exports [1] (Chen et al., 2022). This does not bias our analysis however, since it would mean that for those trade flows this type of benign gap would not appear. If we assume that the incidence of customs fraud is spread

[1] Among large economies, Australia, Brazil, Canada, Mexico, and South Africa are the only countries to use FOB-type valuations for imported goods. On the export side, only six countries (Algeria, Antigua and Barbuda, Burkina Faso, Cambodia, Cape Verde, and Peru) have used non-FOB values for exports, though in each case this was a temporary change to CIF valuation which was reversed after one or two years (Chen et al., 2022).

randomly with respect to the distance of the import country from Sweden, then we would expect that gaps would be more negative (i.e. larger in absolute value) for trade with countries which are further away, since it would cost more to ship goods there. We see indicatively (in Figure A1 in Appendix A) however, that there is instead a slight positive relationship between the outcome variable $ln(Exports) - ln(Imports)$ and the distance between Stockholm and an import country's capital, suggesting that the gaps are likely attributable to other sources.

Separately, since our data is available on a yearly basis, it is possible that due to delays in transit, an export is recorded in the Swedish statistics in one year, and in the import country's statistics in the next. Countries also collect trade statistics in their local currencies, which are then converted to US Dollars by the UN as the COMTRADE database is compiled. Depending on when the statistics are submitted, analysed, and published, differences in foreign exchange rates could also lead to some non-fraud related gaps (Hamanaka, 2011). As above however, these discrepancies are likely to be randomly distributed and therefore while they add noise to our estimates, do not present a threat to identification.

2.3.2 Smuggling

Smuggling, in its most classic sense, involves completely unreported imports or exports. Since we are comparing both sides of a trade flow, if a given shipment is reported neither on export from Sweden nor on import into the destination country, it would not appear in our data. We view however that there is a relatively low risk of this happening, since controls of exported goods are conducted at EU ports. Also, since duties are not charged on export from Sweden, exporters have little incentive to misstate their exports (or not report them at all). Further, in order to not appear in the import country statistics as well, there would have to be a coordinated effort by traders on both sides of the trade flow, which is unlikely on a systematic scale. Smuggling on the import side is more likely, however. If a hypothetical trade flow was properly reported on export from Sweden, but unreported on import, it would also not appear in our dataset. We argue that this would not materially affect our analysis, since the dataset is aggregated on the country - commodity code - year level. As such, individual transactions are not identifiable our dataset, unless there was only one export from Sweden to a given country in a given year of a given product. In all other cases, the dataset would capture all of the other Swedish exports to a country and record a large gap for the smuggled export, validating our approach.

2.3.3 Misclassification

Misclassification is the classification of goods under incorrect or non-applicable HS codes. This could either be through genuine errors on the part of traders or customs authorities, or due to deliberate falsification of documents to conceal the true nature of goods being traded (Hamanaka, 2011). Traders might engage in misclassification to more easily import goods which would otherwise be controlled, or to pass them off as goods which attract lower duties. For example, Afontsev (2012) writes that this was a particularly serious problem in Russia in the late 1990s,

with chicken, which attracted a duty rate of 25%, often being replaced on import declarations with turkey, which attracted duties at a more modest 5%. While in some cases misclassification is obvious, differences between HS codes especially at the HS6 level of granularity can be minute, allowing for tariff rate differentials to be exploited. In terms of our analysis, we would capture such misclassification as large gaps, since a given shipment would be present on one side of the trade flow only (similar to a smuggled shipment as discussed above). Because of this and other related issues discussed in Section 5.3.1, we choose to primarily use the HS4 level for our analysis.

2.3.4 Under- and overvaluation

The main form of tariff evasion that we focus on is the purposeful undervaluation of imported goods by importers. Importers declare values below the true commercial value of their goods in order to benefit from lower import taxes, which are most often calculated based on the declared value. In WTO member countries, customs valuation rules are governed by Article VII of the General Agreement on Tariffs and Trade (World Trade Organisation, 1994; Lyons, 2018), which sets out the methods which customs authorities should use to determine correct values of imported goods. While undervaluation also occurs in the EU, usually through the falsification of commercial invoices (Perotta et al., 2020), this legal framework allows the customs authorities of EU Member-States to act decisively when undervaluation is suspected (European Commission, 2022). In import countries where customs authorities are less capable, there is a higher risk that importers are more able to arbitrarily declare goods as having low values (Cariolle et al., 2019). We argue that this practice is widespread, especially in developing countries, and that it significantly contributes to the trade gaps we see in our analysis.

We also briefly discuss the possibility of overvaluation. This could occur in situations where importers face high taxes on corporate profits in import countries and therefore use inflated payments for imported goods to subsidiaries abroad to move profits to lower-tax jurisdictions (i.e. transfer pricing). For example, Cantens et al. (2012) find that in the case of Cameroonian food imports, there was significant overvaluation, however they also find that this was carried out by a small number of importers who accounted for a large proportion of imports. If transfer pricing occurs, then both the export and import sides of the statistics would be affected. This would mean that the practice would not cause outsized gaps in our dataset, and therefore is not a concern for our analysis.

3 Literature review

The use of mirror trade analysis was pioneered by Bhagwati (1964) in his seminal paper on trade under-invoicing. The use of this methodology increased in popularity as data became richer and more easily available and statistical software became more capable of processing large datasets. Since then, a relatively rich body of literature has emerged linking higher rates

of customs duties and Value Added Tax (VAT) to increased incidence of fraud, which makes intuitive sense since the reward for evading taxes in such cases is higher. Fisman and Wei (2004), for example, find this in China - Hong Kong trade.

Separately, Carrère and Grigoriou (2015) run specifications including geographic factors, GDP per capita, FDI inflows, and measures for the efficiency of customs enforcement. They find that both macroeconomic factors and factors which contribute to evasion incentives (such as duty rates and import country corruption levels) help to explain non-random variations in trade gaps. Javorcik and Narciso (2008) include product differentiation in their analysis and find that tariff evasion is more likely to happen for differentiated products (i.e. non-commoditised products without clear reference prices) due to the difficulties in assessing their value. Later, Javorcik and Narciso (2017) again use mirror trade statistics but this time to study the effect of WTO ascension on tariff evasion. They find that for the fifteen countries in their sample, tariff evasion through under-reported goods valuation decreased after joining the WTO. As outlined in Section 2.3.1, due to operational details of how trade statistics are reported, we expect that in most cases, these gaps will be non-zero (i.e. exports not exactly matching imports). Non-zero gaps are an accepted phenomenon in the literature, for example, Hummels and Lugovskyy (2006) show that only around 50% of trade flows exhibit gaps of less than 100%, though wealthier export countries on average have lower gaps than poorer ones.

Mirror statistics research has largely focused on the import rather than the export side. Levin and Widell (2014) study trade between Kenya and Tanzania, comparing the levels of undervaluation fraud and intentional mislabelling (i.e. where importers purposefully declare their goods as commodities for which duties are lower). They find that undervaluation fraud is more common in Tanzania, while mislabelling is more common in Kenya. They also find that in general, the incidence of customs fraud is lower in Kenya despite Kenya having generally worse corruption scores. This suggests that while institutional quality is key to the relationship between importers and the state, there may be quality differentials across state institutions. Chalendard et al. (2019) conduct a study comparing publicly available trade statistics from the same database we use with internal transaction-level customs authority records from Madagascar. Cariolle et al. (2019) carry out a comparable study focusing on Gabon and come to similar conclusions. Using micro-level data, both sets of authors are able to fully understand the pervasiveness of undervaluation fraud and tie specific customs officers to the particular fraud offences they facilitate. Such work show the importance of public-private partnerships in mirror statistics, especially for data acquisition.

Cantens et al. (2012) conduct a similar exercise in Cameroon, and find that when only ten key products are considered in the analysis, undervaluation fraud in the range of 10 - 20% of total Cameroonian customs revenues is uncovered. They further suggest that while mirror analysis is unlikely to precisely estimate the magnitude of fraud occurring, it can still be used as a tool to assist authorities in their anti-fraud efforts. Nitsch (2015) and Nitsch (2017) attempt to use mirror statistics to quantify the revenue impact of undervaluation fraud and trade misinvoicing

more widely on a global scale and further elaborate on potential issues with identification which we address in Section 5.3. To our knowledge, no author has yet conducted mirror analysis specifically focusing on Swedish export statistics.

While we replicate previous results showing the importance of tariff rates, we are also interested in whether the quality and type of government impact the level of fraud observed. In this vein, Rijkers et al. (2017) conducted a related analysis focusing on imports into Tunisia and found significantly higher rates of customs evasion by companies linked to the deposed dictator Zine El Abidine Ben Ali. Bussy (2021) also includes variables for import country corruption and real income per capita, and finds that undervaluation fraud *decreases* as tariff rates increase in situations where penalties for evasion increase with tariffs. This is the case in most low-corruption high-democracy countries. He also finds however that a reduction in undervaluation fraud does not necessarily result in a net decrease in fraud, since importers may just resort to other methods of evading tariffs. Similarly, Berger and Nitsch (2008) looked at imports into five major economies and examined how trade gaps varied with export country corruption levels. Unsurprisingly, they find that in countries with more corrupt institutions, undervaluation fraud occurs more frequently. Further, Jean and Mitaritonna (2010) find that tariff evasion is more pervasive in lower-income countries, which tend to have worse institutions. When controlling for income, they find that a higher rule of law lowers the extent of the evasion. We would assume that gaps are higher in import countries with less democratic forms of government, however, we are also interested in examining the interactions between the type of government and tariff rates.

4 Data

4.1 Trade data

We primarily use data from the United Nations COMTRADE database, where countries report their trade statistics on a yearly basis (United Nations, 1997). This data is based on the import and export declaration data referenced in Section 2.1, which is provided to the UN by national statistical bodies (Statistics Sweden, 2018). The data we selected for our sample ranges from 2000 to 2021 and includes all traded products (coded at the three different levels of HS granularity). The data includes the type of goods (by HS code), their value in USD, weight in kilograms, and the countries of export and import. In order to build our dataset, we query COMTRADE for Swedish exports to all countries for which data is available from 2000 to 2021. We then query the imports from Sweden reported by each of the import countries for the same time period. Joining these two subsets by year, HS code, and trade flow, we assemble a master dataset of approximately 3 million observations on which we conduct our analysis.

From the raw COMTRADE dataset, we make two exclusions. Firstly, following Rijkers et al. (2017), we exclude "small exports" from the sample. We eliminate exports which make up less

than 0.01% of Sweden's exports for a given commodity code - year pair. In practice, this means that for each year, we consider the total exports from Sweden for a given commodity code and drop from the sample any observations which make up only a minute part of the total. This is done because such values are not representative of wider trends in Swedish exports and therefore risk skewing our later regression results. To ensure that we are not eliminating observations in a way which is correlated with the other covariates, we report the pairwise correlations between all key covariates and a dummy variable for small exports in Table 1. None of the pairwise correlations are high enough to be of concern.

Table 1: Small exports: Correlations with key covariates

	Ln(gap)	Duty %	Distance	GDP	GDPpc	Democracy	Corruption	Small exports
Ln(gap)	1							
Duty %	-0.01	1						
Distance	-0.06	0.24	1					
GDP	0	0	0.11	1				
GDPpc	0.06	-0.24	-0.27	0.18	1			
Democracy	0.04	-0.27	-0.32	0.03	0.5	1		
Corruption	-0.05	0.31	0.3	-0.12	-0.67	-0.77	1	
Small exports	-0.26	0.14	0.19	-0.11	-0.22	-0.24	0.28	1

We also exclude imports and exports attributed to chapter 99 of the HS. This is because this chapter is not officially defined in the HS by the WTO and is used differently in different countries (Berger & Nitsch, 2008). For example in the EU, chapter 99 is used for some low-value intra-EU trade, movements of personal property, deliveries to ships and aircraft, and deliveries to offshore installations such as oil rigs (European Commission, 1987). In Indonesia, the government proposed in 2018 to use chapter 99 for imports of digital goods such as software (Wahyudi & Firdiansyah, 2019). Conversely, in Mozambique, the chapter is used for "other products from other countries", i.e. commodities not captured elsewhere in the trade statistics (Alfieri & Cirera, 2007). Also, as explained in Dahle et al. (1998), countries occasionally "confidentialise" or mask their export statistics for national security or competition reasons. When this happens, the exports are still reported, but they are attributed to chapter 99. Since the exports which are masked are not consistent over time or across countries, we are unable to accurately re-attribute the exports to their correct commodity codes. Therefore, out of prudence, we eliminate these.

4.2 Tariff data

4.2.1 Source and format

We also incorporate tariff data for all WTO member countries for the same period as the COMTRADE data (World Trade Organisation, 2000). We add data from the World Bank World Integrated Tariff Solution (WITS) for non-WTO members (World Bank, 2001)[2]. The data cover 178 countries in total, though not all countries report their tariffs every year. The

[2]The non-WTO countries for which we also have trade data are: Azerbaijan, Bhutan, Bosnia and Herzegovina, Equatorial Guinea, Eritrea, Ethiopia, Lebanon, Libya, Montenegro, Serbia, Sudan, Syria, Turkmenistan, and Uzbekistan.

data itself comprises the tariff rate (charged as a percentage of the declared value) that an import of goods from abroad would face in a particular country, for each commodity code and year. Tariffs are levied on a commodity-code basis, usually at very high levels of granularity (i.e. HS6 or higher). In order to conduct our analysis, we use tariff rates at the HS2, HS4, or HS6 levels. Therefore, it is necessary to aggregate the data. This is done by taking the simple average of all tariff lines "beneath" a certain commodity code. For example, the reported tariff in the data for HS4 commodity code 0101 will be the simple average of the tariffs levied on all HS6 codes beginning with "0101". Tariff rates change over time and can be different for each type of good. In total, we have data for $\approx$ 20 million tariff lines (i.e. country - year - commodity code pairs).

4.2.2 Free Trade Agreement (FTA) usage

While the WTO data mentioned in Section 4.2.1 reports the applicable tariff rate for any given product, it does not include information on Free Trade Agreements (FTAs). These are treaties through which countries agree to lower or eliminate customs duties on trade between them. It is clear therefore that if FTAs are not considered and instead only the raw tariff data is used, this may jeopardise the validity of our results. We therefore also include data from the WTO Regional Trade Agreements database (World Trade Organisation, 2009) and make a number of assumptions about FTA usage on Swedish exports.

When EU Member-States are included, Sweden currently has such agreements with ninety-nine countries (for the full list, see Appendix F). Since our analysis is conducted on a yearly basis, we assume that a trade agreement came into force on the first day of the year in which it was signed. Further, some FTAs exclude certain chapters of the HS or introduce reduced duties for different parts of the HS at different periods of the implementation process. Due to the complexity of analysing each FTA separately, we assume that every FTA covers the entire HS (i.e. all products), and that it eliminates all customs duties immediately upon coming into force. In order to counter these somewhat strong assumptions, we introduce a number of variables for the applicable tariff, considering different rates of FTA usage. This accounts for both the fact that not all products are included in every FTA in reality, and also that not all importers claim FTAs. This could happen if importers are not aware that their shipments qualify for exemption from duties under an FTA, or if they are not be willing to take on the administrative burden of proving their goods' eligibility, preferring instead to pay the duties on them.

We show in Appendix D that the results are not sensitive to different FTA use assumptions. Out of prudence, we therefore proceed with the assumption that all FTAs are utilised 100% of the time that they are applicable. While this is likely an overestimation, it ensures that we do not overstate the importance of our main covariate, the customs duty rate. Estimates provided for the effect of customs duties on the trade gap therefore constitute a lower bound of the potential effect.

4.3 Other covariates

For measures of democracy and corruption, we use the Varieties of Democracy (VDEM) dataset (Coppedge et al., 2022a). Specifically, we use the *v2x_libdem* and *v2x_pubcorr* variables, respectively. Each is an index with values between 0 and 1, however with different directions. *v2x_libdem* measures the extent to which liberal democracy is achieved, with 0 signifying no democracy at all and 1 signifying full, perfect democracy. This measure is a composite of two indices, one which measures electoral democracy through scores for freedom of association, freedom of expression, the quality of the electoral process, the pervasiveness of the electoral system in selecting a country's leaders, and the extent to which suffrage is universal, and the other which includes measures of judicial and legislative constraints on executive power, as well as useful measures for state capacity. Conversely, *v2x_pubcorr* is an aggregate of measures for the average level of bribery and kickbacks in the public sector, as well as measures for theft of state property by public sector officials. The index being coded 0 when there is no corruption at all, and 1 when corruption is at its maximum level (Coppedge et al., 2022b). We also use data from the World Bank on GDP and GDP per capita. Finally, data on distances between Stockholm and import country capitals is from the *tradestatistics* R package (Vargas, 2022).

5 Empirical strategy

5.1 Hypotheses

In this section, we lay out the three hypotheses that are the basis for our analysis and which we aim to confirm empirically.

The first hypothesis is that countries further away from Sweden will have more negative gaps. The reasoning behind this is as outlined in 2.3.1. It relies on the fact that imports are reported on a CIF basis, i.e. with the costs for transportation and insurance included in the total value of the goods, while export statistics are calculated on the FOB basis, meaning that they include only the value of the goods themselves. If we assume that the incidence of customs fraud is spread randomly with respect to distance from Sweden, then we would expect an inverse relationship between the outcome variable and the distance. The rationale here is that as distance increases, transportation costs would also increase and so will the reported import value, and the bigger the import value relative to the export one, the more negative the trade gap will be. This has been shown by Hummels and Lugovskyy (2006) in the global context, but we aim to confirm its applicability in the context of Swedish exports.

Hypothesis two is that an increase in import duties will lead to an increase in the trade gap. Since import taxes are in general levied on the value of the imported goods, importers are incentivised to understate this value and therefore pay lower taxes. As mentioned above, this effect has been consistently shown in the literature, for example in the seminal paper by Fisman and Wei (2004) analysing Chinese imports, and also in Levin and Widell (2014), which looks into

imports into Tanzania and Kenya. However, to our knowledge, no researcher has yet conducted this analysis specifically on Swedish exports. Countries with well-developed and competent tax authorities are usually able to stop or at least reduce the incidence of this type of customs evasion, however in countries where this is not the case, undervaluation fraud is a relatively easy and low-risk way for importers to evade taxes. In some countries, this may also take place with the complicity of corrupt customs officials, as uncovered by Chalendard et al. (2019).

Our third hypothesis is that gaps will be higher in countries with weaker institutions, as measured by lower democracy scores and higher levels of endemic corruption. We further believe that the effect of increased import duties will be stronger for countries with weaker democracies. Due to worse institutions such as the lack of strong rule of law and high levels of corruption, we argue that evasion is easier in weakly- and non-democratic countries. If importers do not fear legal sanctions from their peers or the state, they are more likely to act in their own self-interest and lower their tax bills. We also run an empirical specification which controls for the level of corruption in a country, allowing us to compare countries which have the same level of "background" corruption, but which differ in the strength of their democracy. We argue that in this case, more democratic countries will still have lower gaps than less democratic ones.

5.2 Specifications

5.2.1 Outcome variable

The outcome variable of interest in this study is the gap between the officially reported gross export flows and the officially reported gross import flows. Following the example of Javorcik and Narciso (2008) and Bussy (2020), we use log differences in value as the main outcome variable. The gap is calculated as:

$$TradeGap_{ijk} = log(\frac{Export_{ijk}}{Import_{ijk}}) = log(Export_{ijk}) - log(Import_{ijk}) \quad (1)$$

Where *Export* is the export country figure, *Import* is the import country figure, i is the import country, j is the year, and k is the HS code.

5.2.2 Baseline specification

$$y_{ijk} = \alpha_{ijk} + \beta_1 Tax_{ijk} + \beta_2 Democracy_{ij} + \beta_3 Corruption_{ij} + \beta_4 Distance_i + \beta_6 GDP_{ij} + \beta_7 GDPpc_{ij} + \epsilon_{ijk} \quad (2)$$

Where: y is the outcome variable, *Tax* is the average duty percentage for commodity code k in import country i and year j (assuming that all applicable FTAs are fully utilised), *Democracy* and *Corruption* are indexes from the VDEM dataset, with full democracy coded as 1 and the least corruption coded as 0, again for the import country and year, *Distance* measures the

distance from Stockholm to the capital city of each trading partner, *GDP* is the import country GDP in levels in USD, and *GDPpc* is the import country GDP per capita.

We run various specifications based on Equation 2 which are presented in Table 3 as specifications (1) to (6). If our first hypothesis (i.e. that the log trade gap should be lower in countries further away from Sweden) is correct, then β_4 should be significant and negative. Also, if our second hypothesis (i.e. that the log trade gap, and therefore undervaluation fraud increase with higher duty rates) is correct, β_1 should also be positive. Further, in order to support hypothesis three (i.e. that gaps will be higher in weaker democracies), β_2 should be negative.

5.2.3 Fixed Effects specifications

$$y_{ijk} = \alpha_{ijk} + \beta Tax_{ijk} + X + \zeta_i + \gamma_j + \epsilon_{ijk} \tag{3}$$

Where: ζ_i is import country fixed effects and γ_j is year fixed effects.

Equation 3 is the same as Equation 2, except for the addition of the fixed effects. This enables us to test the same hypotheses, but with the added controls for variation across time and import countries. Specifically, the inclusion of year fixed effects allows us to control for any time trends in the gaps. Country fixed effects address some concerns we had about omitted variable bias, as explained in Section 5.3.2. We report the results from Equation 3 as specifications (7) to (9) in Table 4.

5.2.4 Interaction terms

$$\begin{aligned} y_{ijk} = \alpha_{ijk} + \beta_1 Tax_{ijk} + \\ \beta_2 Tax * LowDem_{ijk} + \beta_3 Tax * MedDem_{ijk} + \beta_4 LowDem_{ij} + \beta_5 MedDem_{ij} + \\ \beta_6 Tax * MedCorr_{ijk} + \beta_7 Tax * HighCorr_{ijk} + \beta_8 MedCorr_{ij} + \beta_9 HighCorr_{ij} + \\ X + \zeta_i + \gamma_j + \epsilon_{ijk} \end{aligned} \tag{4}$$

Where the *Low*, *Med*, and *High* prefixes describe the tertiles of respective variables. These tertiles are recalculated for both democracy and corruption scores on a yearly basis to cmaintain balance in these variables over time.

We report the results from Equation 4 as specifications (10) to (12) in Table 5. In order to properly investigate whether our third hypothesis (i.e. that the impact of higher duty rates on higher incidence of undervaluation fraud is stronger in less democratic countries) is true, we use interaction terms. If β_2 and β_3 are positive and significantly different from zero, we would argue that indeed the effect of higher tariff rates on increasing trade gaps is more pronounced in weakly-democratic countries. Further, we expect that β_6 and β_7 will also be positive, since increases in the tariff rate are logically more likely to cause evasion in countries which are highly corrupt, and where evasion is therefore easier.

5.3 Potential threats to identification

5.3.1 HS code mismatches

Misclassification occurs when goods imported under a given commodity code are exported under another. Such a situation would cause large gaps in both commodity codes as only one side of the trade flow would be present in the data. This is a larger risk at higher levels of granularity. Similar to the misclassification issue discussed in 2.3.3 is the problem of "concordance". This is when HS codes from different versions of the HS are compared and normalised. As discussed above in Section 2.2, the HS is largely stable over time. However, the WTO updates it every five years in order to ensure that new products are properly described and that the HS as a whole continues to be fit for purpose. HS codes can be updated and either brought into service, retired, or changed to cover different goods. When reporting their trade statistics and tariff rates, countries can choose which version of the HS to use, and usually update their use of the HS a few years after the new HS is adopted by the WTO. However, there is a risk that developing countries are more likely to be late in updating the version of the HS they use, whereas wealthier countries make the change to newer versions of the HS faster (Cantens et al., 2012). Our data show that mismatches are indeed more prevalent in countries with low Human Development Index (HDI) scores, which may introduce bias in our results. This is clearly shown in Figure A2 in Appendix A.

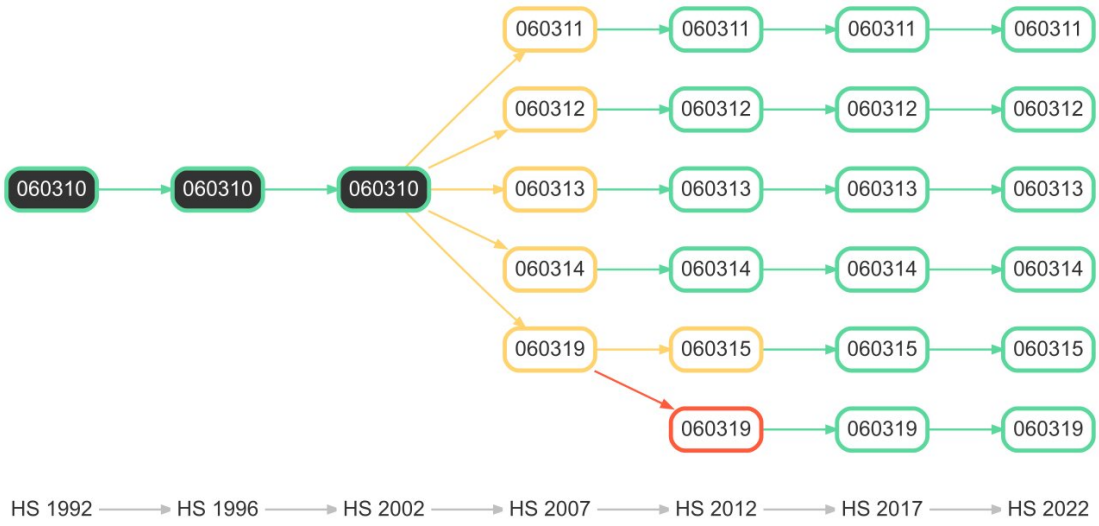

Figure 2: Changes to the HS for HS code 060310. Note that codes in green are unchanged from previous versions, orange are new, and red are recycled from previous versions. Source: World Trade Organisation (2023).

If an import country and Sweden use different versions of the HS, there is a risk that we would record a large gap where this is simply due to a concordance issue. Reconciling HS codes from various versions is not feasible in this case due to the vast number of HS codes and country combinations. An example of changes to the HS is provided in Figure 2, showing the progression of HS codes used for 060310 "Fresh cut flowers and flower buds, for bouquets or for ornamental

purposes". While there was a significant change in the 2007 version of the HS, before and after this the HS codes are largely stable, and in all cases the HS4 code (i.e. 0603) is maintained. We argue that it is relatively unlikely that products change HS4 classification due to concordance issues, even if their HS6 classification changes. We therefore rerun our analyses at the HS2, HS4, and HS6 levels of granularity and compare the results. Since we cannot exclude the possibility that the results at the HS6 level are driven by misclassification or concordance issues and do not reflect true effects, we diverge from the literature (which usually focuses on the HS6 level) and decide to run specifications using the HS4 data instead. This is discussed in more detail in Section 6.5.1.

5.3.2 Omitted variable bias and measurement error

Omitted variable bias is a significant threat to our identification strategy. If some factors not included in our model are correlated with both our covariates and the outcome variable, our estimates are likely to be biased. For example, we assume that in the absence of fraud, trade gaps vary randomly. If instead there is a country characteristic which is likely to affect both the trade gaps and the tariff rates, then our identifying assumptions could be violated.

Similarly, measurement error could impact the biasedness of our estimates. For example, we assume that time mismatches (i.e. where different sides of the trade flow are recorded in successive years due to shipping delays), differences in classification rules, and exchange rates are random across countries. If this is not the case, for example if certain types of import countries are consistently late in reporting imports, or if they have policies for valuing shipping and insurance costs which differ from other countries, then this could be a potential threat to identification (Bussy, 2020). We control for this by using country fixed effects as well as time fixed effects, and including controls accepted in the mirror statistics literature such as GDP and GDP per capita.

5.3.3 Reverse causality

While it is theoretically possible that tariff rates are set in response to the levels of evasion for given products, we view that this is relatively unlikely. As noted by Bussy (2021) this is much more difficult for countries in the WTO, where tariffs are set based on multilateral negotiations, and countries cannot freely change them as a response to evasion. Since our sample consists almost exclusively of countries that are part of the WTO, reverse causality is less of a concern. Swedish exports to the non-WTO countries mentioned in Section 4.2.2 comprise only 0.17%[3] of the total, and are therefore negligible. Separately, low- and middle-income countries are often constrained by commitments to international organisations such as the World Bank and the International Monetary Fund (IMF) to liberalise trade taxes, presenting a further obstacle to arbitrarily raising tariffs in the face of rising undervaluation fraud (Biersteker, 1990).

[3]The sum of all Swedish exports by value to WTO countries from 2000 to 2021 is ≈ USD 2.958T, whereas for non-WTO countries the figure is ≈ USD 5.1B.

6 Results and discussion

6.1 Descriptive statistics

Table 2: Summary Statistics

Variable	N	Mean	Std. Dev.	Min	Pctl. 25	Pctl. 50	Pctl. 75	Max
Ratio EXP/IMP USD	655934	53.685	4939.183	0	0.606	1.037	2.222	2746936.571
Log Gap USD	655934	0.188	1.682	-11.946	-0.501	0.036	0.798	14.826
Average Duty Percentage	579574	3.173	10.216	0	0	0	5	2915.143
Democracy (1 = best)	655109	0.637	0.26	0.006	0.468	0.772	0.826	0.895
Corruption (0 = best)	655284	0.224	0.255	0.001	0.023	0.106	0.399	0.966

Table 2 shows the summary statistics of the sample. Overall there are 655,934 observations for which we have data on trade value and out of those there is tariff data for 579,574 observations. More than half of the sample has a tariff rate of 0%, which is visible as well in Figure 3, where we plot the density of the duty percentages. This is easily explained by the prevalence of Swedish exports to the EU (all attracting duties at 0%) and of wider global trade liberalisation trends (for example, see Figure A3 in Appendix A for the downward trend in customs duties on Swedish exports over time).

More than half of the observations have an export-to-import ratio larger than 1, suggesting a strong tendency to under-report imports, which is most likely a mix of benign and systemic factors. The mean log gap in value is 0.188 (i.e. the outcome variable calculated as in Equation 1), meaning that on average reported imports are 18.8% smaller than exports. Considering that imports are reported including the cost of freight and insurance, whereas exports are not, an average under-reporting of 18.8% further reinforces our hypothesis that these statistical gaps are more than just honest mistakes. The last two lines show the summary statistics for the democracy and corruption scores of the sample. The minimum democracy score is 0.006, while the 25^{th} percentile is 0.468, which means that the majority of the countries in our sample score highly on the electoral democracy scale. A similar pattern is observed in the corruption scores, where 75% of the observations have a corruption score of 0.399 or less.

This means that our sample is unbalanced with regard to those variables. The lack of balance is due to the way trade data is aggregated. Since we have data on a country-year-product level, countries to which Sweden exports a wider variety of products appear more often in the dataset. A further illustration of this can be found in Figure A5 in Appendix A, which shows that countries with above-average HDI scores make up a large part of Swedish export destinations when only tariff lines and not trade value are considered. We argue that this is not especially concerning for our analysis since our data covers the entire universe of Swedish exports to all countries for the years 2000 - 2021. Our sample is therefore unbalanced because Swedish exports are also inherently unbalanced. Nonetheless, we manage this in a number of ways. Firstly, we introduce tertiles for for the democracy and corruption variables which are calculated separately for each year to alleviate bias concerns. This is explained in more detail in Section 6.4. Also, we

run our main specifications giving equal weights to each country and find that our results are largely unchanged. This is discussed in further Section 6.2.

Figure 4 shows that the median gaps decrease as the HDI of import countries increases and that the biggest outliers tend to be countries with low HDI scores, which are also more likely to have worse institutions and higher corruption. Lastly, Figure 5 decomposes the log median gaps by HS2 chapter, relative to the weighted average for the whole HS. For the majority of the HS2 chapters, the median log gaps lie above 0. The highlighted outliers (i.e. those HS2 chapters with median log trade gaps > 0.5) are chapters 46 (straw products), 61 (knitted clothing), 62 (other clothing), 64 (footwear), 67 (feather products), and 91 (clocks and watches), which are all relatively small and potentially easy-to-smuggle items. All are also differentiated products, lending further support to the argument put forward by Javorcik and Narciso (2008) that importers of these products are more likely to engage in tariff evasion. Figures A5 to A10 in Appendix A further decompose the relevant HS2 chapters into their HS4 sub-chapters to show that while there is a large variation with an HS2 category, in this case the vast majority of HS4 codes exhibit significant trade gaps.

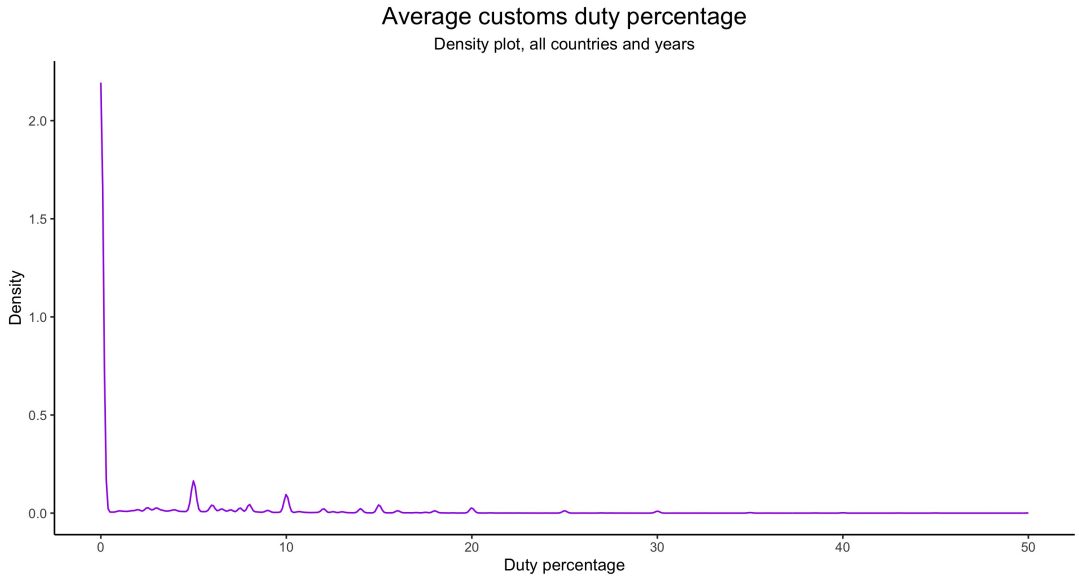

Figure 3: Density of customs duty rates charged on all Swedish exports, 2000 - 2021

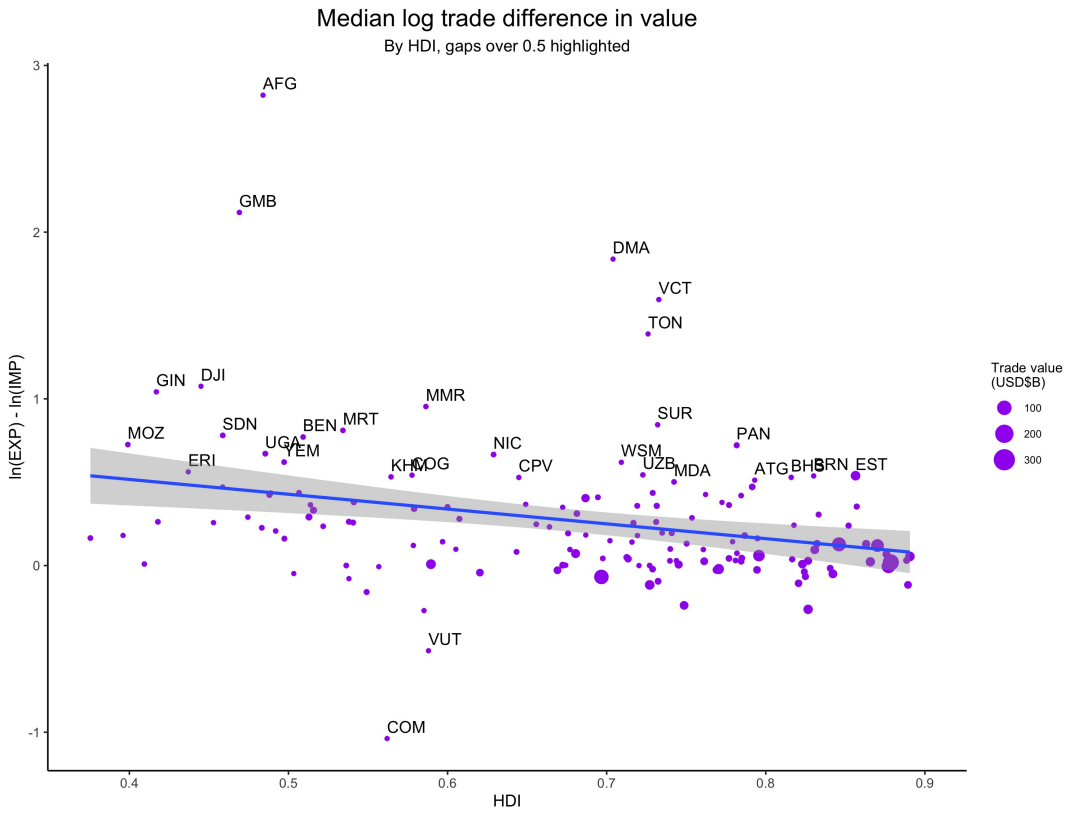

Figure 4: Median trade gap by Human Development Index (HDI), 2000 - 2021

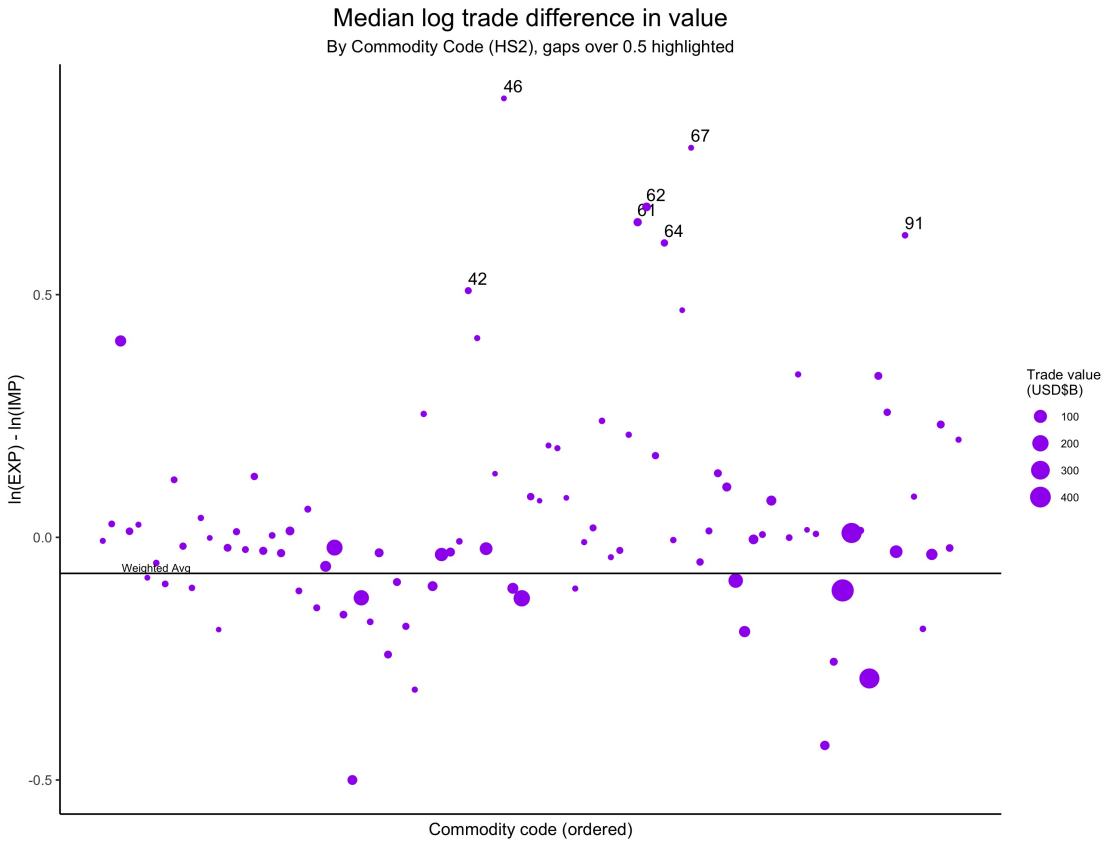

Figure 5: Median trade gap by HS2 chapter, 2000 - 2021

6.2 Baseline specification

Our first six specifications are shown in Table 3. Using the log trade gap as the dependent variable, we see that even when no other covariates are included in the regression, the duty rate is significant and positive. This means that, following the mirror statistics literature and our second hypothesis, we would expect the log trade gap (and therefore also undervaluation fraud) to increase as a result of duty rates increasing. Specifically, we find in specification (1) that a one percentage point increase in the duty rate leads to a 0.5% increase in the the log trade gap, an effect which is significant at the 1% level. This relationship and its significance is robust to the inclusion of various covariates and fixed effects.

When introducing the democracy variable on its own in specification (2), its coefficient is highly significant (at the 1% level) and negative. Since the variable is coded from zero to one, where zero is the worst possible democracy and one the best, this result makes sense. In essence, if a product attracting the same duty rate is imported into two countries, we would expect that fraud is more likely to happen in the less democratic country. Similarly, specification (3) shows the results for the corruption variable in isolation. In this case, it is coded so that a value of zero indicates no corruption and a value of one is the most possible corruption. The coefficient is positive and highly significant (at the 1% level). Therefore, as countries become more corrupt, the log trade gap and therefore undervaluation fraud increase, *ceteris paribus*. As with the democracy variable, this follows our expectations. When controlling for both democracy and corruption in column (4), the significance and direction of the effect of democracy on the trade gaps remains unchanged. Corruption does lose significance but remains generally positive. This is not concerning however, since in this regression we still do not control for the most benign determinant of the trade gaps, i.e. distance.

When we include the distance of the import country's capital from Stockholm into the regression in specification (5), we see that the coefficient is significant at the 1% level, and that it is negative. While the coefficient seems small and potentially economically insignificant, it is important to note that the distance variable is coded in kilometers. Therefore, on average for every thousand kilometers further away an import country's capital is from Stockholm, we expect the log trade gap to be 1% lower. This is logical due to how the log trade gap variable is constructed and the discrepancy between CIF and FOB trade statistics as explained in Section 2.3.1 and stated in our first hypothesis. Due to the CIF method's near-omnipresence in import country statistics, this variable is an important control and it is encouraging that it has the expected sign.

Next, in specification (6), we also include controls for import country GDP in levels, and import country GDP per capita. All coefficients are significant at the 1% level. For GDP and GDP per capita, we find that the coefficients are extremely small (≈ 0), though they are very precisely estimated. Although they have no effect, we continue to use them as controls throughout the later specifications out of prudence. Reassuringly, these naive OLS results remain very consistent

Table 3: Sweden - Regressions with different covariates

	Dependent variable:					
	lnGap_USD					
	(1)	(2)	(3)	(4)	(5)	(6)
Duty Percentage	0.005***	0.004***	0.004***	0.004***	0.006***	0.007***
	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.001)	(0.001)
Democracy (1 = best)		−0.081***		−0.067***	−0.105***	−0.090***
		(0.010)		(0.014)	(0.014)	(0.014)
Corruption (0 = best)			0.070***	0.019	0.028*	0.234***
			(0.010)	(0.015)	(0.015)	(0.017)
Distance to capital (km)					−0.00001***	−0.00001***
					(0.00000)	(0.00000)
GDP						−0.000***
						(0.000)
GDP per capita						0.00000***
						(0.00000)
Constant	0.166***	0.221***	0.154***	0.209***	0.273***	0.116***
	(0.002)	(0.007)	(0.003)	(0.012)	(0.012)	(0.014)
Observations	613,481	612,761	612,932	612,761	612,761	612,761
R^2	0.0003	0.0004	0.0004	0.0004	0.001	0.003
Adjusted R^2	0.0003	0.0004	0.0004	0.0004	0.001	0.003

Notes: Data are for Swedish exports to all countries, from 2000 - 2021, aggregated at the HS4 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of versions of Specification 2. The outcome variable is ln(Exports) - ln(Imports). The duty percentage refers to the amount of customs duty levied on imports of Swedish goods into a given import country. As explained in Section 4.2.2, we assume that any applicable Free Trade Agreements are fully utilised. The democracy variable, from the VDEM dataset, is a composite of two indices, one which measures electoral democracy through scores for freedom of association, freedom of expression, the quality of the electoral process, the pervasiveness of the electoral system in selecting a country's leaders, and the extent to which suffrage is universal, and the other which includes measures of judicial and legislative constraints on executive power, as well as useful measures for state capacity. The corruption variable is an aggregate of measures for the average level of bribery in the public sector as well as measures for theft of state property by public sector officials. Both the democracy and corruption variables are from the VDEM dataset. The distance variable records the distance between Stockholm and the import country's capital. The GDP and GDP per capita relate to the import country. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

and follow our expectations in terms of the direction of the effect as well as it's significance and lend support to our third hypothesis that gaps are likely to be higher in countries with weaker institutions.

It is important to mention, as explained in 6.1, that the sample is fairly unbalanced in terms of democracy and corruption, due to high-democracy low-corruption countries being the main destinations for Swedish exports. To test whether or not the results we observe are driven by certain countries in our sample, we re-run the specification with equal weights for each import country. The results are presented in Table E1 in Appendix E, and are broadly identical to the non-weighted ones, ruling out that the effect is caused by a small number of non-representative counties. It could also be however that these results are driven by unobserved country characteristics, which are not controlled for with the limited selection of controls we have in specifications (1) to (6). Ultimately these are relatively naive OLS estimations, which do not adequately take our concerns about threats to identification (especially omitted variable bias) into consideration. In order to address these, we continue with the fixed effects specifications.

6.3 Fixed Effects specifications

For convenience, we report specification (6) in Table 4, which is identical to specification (6) in Table 3. In this specification, no fixed effects are included. In specifications (7), (8), and (9), we first include import country fixed effects, then year fixed effects, and then both. Distance is not reported in columns (7) and (8), as it is absorbed by the country fixed effects.

We find that the coefficient on duty percentage continues to be positive and highly significant (at the 1% level) even when both country and year fixed effects are added to the specification. This is encouraging as it provides further support for our second hypothesis that as the duty rate increases, so does the risk and incidence of undervaluation fraud. For the democracy variable, we see that the direction and significance of the estimate remain robust to the addition of both types of fixed effects in columns (7) and (8). Despite removing a large amount of variation by using two-way fixed effects in specification (9), the estimate is still negative and significant at the 10% level.

The results for endemic corruption mirror those for democracy. In each specification, the coefficient on the variable is positive as expected, at the 5% level in specification (7), and the 1% level in specifications (8) and (9). In specification (9), even when country and year fixed effects are included in the regression, the coefficient remains quite large, showing that on average, if a country went from the lowest to highest possible levels of corruption, the log trade gap would increase by 23.2%. These results are encouraging since they alleviate, at least somewhat, our identification concerns outlined in Section 5.3.

Table 4: Sweden - Fixed Effects specifications

	Dependent variable:			
	lnGap_USD			
	(6)	(7)	(8)	(9)
Duty Percentage	0.007***	0.009***	0.007***	0.009***
	(0.001)	(0.001)	(0.001)	(0.001)
Democracy (1 = best)	−0.090***	−0.292***	−0.077***	−0.100*
	(0.014)	(0.048)	(0.014)	(0.052)
Corruption (0 = best)	0.234***	0.116**	0.246***	0.232***
	(0.017)	(0.058)	(0.018)	(0.060)
Distance to capital (km)	−0.00001***		−0.00001***	
	(0.00000)		(0.00000)	
Constant	0.116***			
	(0.014)			
Country Fixed Effects	No	Yes	No	Yes
Year Fixed Effects	No	No	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	612,761	612,761	612,761	612,761
R^2	0.003	0.038	0.014	0.038
Adjusted R^2	0.003	0.037	0.014	0.038

Notes: Data are for Swedish exports to all countries, from 2000 - 2021, aggregated at the HS4 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of Specifications 2 and 3. The outcome variable is ln(Exports) - ln(Imports). Column (6) presents our preferred version of Specification 2, i.e. with all the covariates included, but with no Fixed Effects. Columns (7) to (9) add Country FE, Year FE, and then both. The variables are as described in the notes to Table 3. The control variables not shown in the table are the GDP and GDP per capita. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

6.4 Interaction terms

In order to investigate our third hypothesis, that the effect of increased import duties on undervaluation fraud will be stronger for countries with weaker democracies, we rerun the regressions, but this time using interaction terms between the duty rate and democracy and corruption tertiles. Specification (10) includes only democracy interaction terms, specification (11) has only those for corruption, and specification (12) has both. All specifications also include country and year fixed effects, as well as controls.

As before, in all specifications the coefficient on the duty percentage remains positive and significant at the 1% level. Taking specification (10) first, we see that there is a much larger impact of duty rates on the log trade gap in countries in both the low and medium democracy tertiles, as compared to countries in the highest democracy tertile. In countries in the lowest democracy tertile, an increase of one percentage point in the duty rate will on average lead to a 1.3% increase in the log trade gap. In countries in the medium democracy tertile, the figure is 0.9%. This is compared to 0.6% for countries in the high democracy tertile. The coefficients for the low democracy and medium democracy interaction terms are also significantly different from each other at the 1% level. These results confirm our third hypothesis, since weaker democracies exhibit more than twice as large an effect on fraud for duty increases as strong democracies. This is likely because penalties for customs evasion are usually lower in weaker democracies (Bussy, 2021), corruption levels are also higher, and the capacity of the state to enforce customs regulation is low, all facilitating undervaluation fraud. For import countries in the low and medium democracy tertiles, we see that on average, there are respectively 6.6 % and 7.5% higher log trade gaps than in high democracy countries, respectively. This lends further support to our third hypothesis in that democracy levels in import countries are clearly an important determining factor in the level of gaps, and therefore also the incidence of undervaluation fraud.

When only the corruption interaction terms are included in specification (11), both the interaction terms of the duty percentage with the high and medium corruption tertiles are positive and significant. In low corruption countries, a one percentage point increase in the duty rate leads to a 0.5% increase in the log trade gaps. In medium corruption tertile countries, it leads to a 1.2% increase, whereas in high corruption countries it leads to a 0.9% increase. All of these effects except the differential effect in high corruption countries are significant at the 1% level. The high corruption interaction term is instead significant at the 10% level. Further, the medium and high corruption interaction terms are significantly different from each other at the 1% level. This is what we would expect, since it shows that on average, the log trade gaps caused by duty increases are higher in high corruption countries, where tariff evasion is more endemic and therefore easier for importers. It is puzzling, however, that the effect is more pronounced in medium corruption tertile countries than in those with high corruption. We view that both the low magnitude of the coefficient and its low significance are attributable to the same unbalancedness mentioned above. Whereas most of the countries in our sample are clustered around high levels of democracy, they also largely have low levels of corruption.

Table 5: Sweden - Regressions with interaction terms

	Dependent variable:		
	lnGap_USD		
	(10)	(11)	(12)
Duty Percentage	0.006***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)
Duty x low democracy	0.007***		0.005**
	(0.002)		(0.002)
Duty x medium democracy	0.003**		−0.001
	(0.001)		(0.002)
Low democracy	0.066***		0.072***
	(0.023)		(0.024)
Medium democracy	0.075***		0.084***
	(0.015)		(0.016)
Duty x medium corruption		0.007***	0.006***
		(0.001)	(0.002)
Duty x high corruption		0.004*	0.0002
		(0.002)	(0.003)
Medium corruption		0.005	−0.006
		(0.018)	(0.018)
High corruption		0.018	0.005
		(0.025)	(0.026)
Country Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Controls	Yes	Yes	Yes
p-value: duty*lowDem = duty*medDem	<0.01		<0.01
p-value: duty*medCorr = duty*highCorr		<0.01	<0.01
Observations	612,761	612,932	612,761
R^2	0.038	0.038	0.038
Adjusted R^2	0.038	0.038	0.038

Notes: Data are for Swedish exports to all countries, from 2000 - 2021, aggregated at the HS4 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of versions of Specification 4 with different interaction terms included. The outcome variable is ln(Exports) - ln(Imports). Column (12), our preferred specification, includes interaction terms between the duty percentage and tertiles of both the democracy and corruption scores. Tertiles for the democracy and corruption variables are recalculated on a yearly basis. The variables are as described in the notes to Table 3. The control variables not shown in the table are the GDP and GDP per capita. The p-values shown in the bottom of the table refer to a two-sided t-test verifying the equality of the coefficients on different interaction terms. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

We see that specification (12) broadly maintains the results from specifications (10) and (11), however a number of variables lose significance. Even so, our main effect is robust to the inclusion of both sets of interaction terms. As before, the duty rate coefficient is positive and highly significant. We find that in high democracy, low corruption countries, a one percentage point increase in the duty rate leads to, on average, a 0.5% increase in the log trade gap. While there is no significantly different effect in medium democracy countries, in low democracy countries, a similar increase in the duty percentage will on average lead to a 1% increase in the log trade gap. In other words, for two countries which fall in the same corruption tertile, an increase in the customs duty will likely have a more detrimental effect on the lower democracy one. As in specficiation (10), both the low democracy tertile coefficient and the medium democracy tertile coefficient are large, positive, and significant at the 1% level. For the corruption interactions with the duty rate, the high corruption tertile has no significant effects, but the medium corruption tertile does. In the relevant import countries, a one percentage point increase in the duty rate leads to, on average, a 1% increase in the log trade gap, compared to a 0.5% increase in low corruption import countries. For both the democracy and corruption interaction terms, the respective coefficients are different from each other at the 1% level.

6.5 Robustness checks

6.5.1 Results at different levels of HS granularity

In Table C3, we report our main interaction specifications run at the three levels of HS granularity. Full results covering all of the specifications are reported in Appendix C.

For the naive OLS estimations like that of specification (6) of Table 3, we find that the results hold across the different levels of HS granularity. Specifically, they are broadly similar in terms of the sign, magnitude, and level of significance of the coefficients. However, when adding fixed effects as in Table 5 we find that the HS6 level presents some surprising results. For example, in Table C3 the interaction term for duty and high corruption is highly significant and negative, indicating that trade gaps *decrease* as duty rates increase in countries with high levels of corruption. This goes against the accepted literature and also common sense.

Following our arguments on misclassification and concordance outlined in Section 5.3.1, we cannot rule out that the results at the HS6 level are driven by bad-quality data where gaps are recorded due to HS version mismatches or intentional misclassification instead of undervaluation fraud. Separately, when considering the HS2 level of granularity, we find that in Table C3 almost all of our covariates cease to be significant at any of the traditional levels of significance. This is unsurprising, since due to the aggregation of the data into broad HS2 categories, there is much less variation to explain, reinforcing our choice to conduct the analysis at the HS4 level.

Table 6: Regressions with interaction terms at different HS levels

	Dependent variable:		
	lnGap_USD		
	(HS6)	(HS4)	(HS2)
Duty Percentage	0.013***	0.005***	0.003**
	(0.001)	(0.001)	(0.002)
Duty x low democracy	−0.001	0.005**	0.010**
	(0.002)	(0.002)	(0.004)
Duty x medium democracy	−0.005***	−0.001	0.008**
	(0.001)	(0.002)	(0.003)
Low democracy	0.147***	0.072***	−0.001
	(0.018)	(0.024)	(0.044)
Medium democracy	0.168***	0.084***	0.008
	(0.012)	(0.016)	(0.030)
Duty x medium corruption	−0.001	0.006***	−0.001
	(0.002)	(0.002)	(0.003)
Duty x high corruption	−0.009***	0.0002	−0.003
	(0.002)	(0.003)	(0.004)
Medium corruption	−0.010	−0.006	0.013
	(0.014)	(0.018)	(0.036)
High corruption	−0.005	0.005	0.010
	(0.020)	(0.026)	(0.049)
Country Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Controls	Yes	Yes	Yes
p-value: duty*lowDem = duty*medDem	<0.01	<0.01	0.2205
p-value: duty*medCorr = duty*highCorr	<0.01	<0.01	0.516
Observations	1,504,167	612,761	93,399
R^2	0.035	0.038	0.050
Adjusted R^2	0.035	0.038	0.048

Notes: Data are for Swedish exports to all countries, from 2000 - 2021, with the data aggregated at the HS2, HS4, and HS6 levels of the harmonised system. Robust standard errors reported in parentheses. This table reports the results of Specification 4 with all interaction terms included. The outcome variable is ln(Exports) - ln(Imports). Tertiles for the democracy and corruption variables are recalculated on a yearly basis. The variables are as described in the notes to Table 3. The control variables not shown in the table are the GDP and GDP per capita. The p-values shown in the bottom of the table refer to a two-sided t-test verifying the equality of the coefficients on different interaction terms. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

6.5.2 Comparison with other countries

While they are not directly comparable to Sweden due to differing economic circumstances and export profiles, we also acquire data for exports from Finland, Denmark, and Norway from 2000 to 2021. We then rerun all of our analyses and present the comparative results in Table G3. Full results for Finland, Denmark, and Norway are reported in Appendices G.2, G.3, and G.4, respectively and additional comparisons can be found in G.1.

Encouragingly, we find that our main results largely hold across the Scandinavian region. In all specifications except one (Norway in Table G3), the coefficient on the duty rate is positive and highly significant. Further, there are significant differential effects for duty rate increases in low democracy import countries for all of the Scandinavian countries except Denmark. Unfortunately, a detailed investigation of each export country is beyond the scope of this thesis, however it presents an interesting opportunity for future research.

6.5.3 Tertile balance

As discussed in Section 6.1, we had some concerns that our sample was unbalanced with regard to corruption and democracy scores and that this would bias our results. While we address these with an import country-weighted regression (as discussed in Section 6.2), using tertiles for the interaction also presents a separate issue. Namely, that we are only able to identify the variation in the log trade gaps caused by countries which *change* either democracy or corruption tertiles in the sample. If this subsample of countries is significantly different from the main sample, then our results would risk losing internal validity. The subsample contains fifty-two countries whose corruption tertile changes, and fifty-eight countries whose democracy tertile changes, out of a total of 170 countries. The full list of countries and the number of times they change tertile is presented in Appendix B.

As shown in Figures B1 and B2 in Appendix B, our concerns are largely unfounded. While the median HDI and log trade gaps are slightly different for the main and subsamples, they are extremely similar. Further, the fitted trendlines for the number of times a given country changes tertile are similar for the corruption and democracy measures.

Table 7: Regressions with interaction terms for different export countries

	Dependent variable:			
	lnGap_USD			
	(SWE)	(FIN)	(NOR)	(DNK)
Duty Percentage	0.005***	0.009***	0.001	0.007***
	(0.001)	(0.001)	(0.001)	(0.002)
Duty x low democracy	0.005**	0.009***	0.008***	−0.004
	(0.002)	(0.003)	(0.002)	(0.004)
Duty x medium democracy	−0.001	0.0001	−0.001	0.004
	(0.002)	(0.002)	(0.002)	(0.003)
Low democracy	0.072***	0.162***	0.020	0.144***
	(0.024)	(0.030)	(0.027)	(0.039)
Medium democracy	0.084***	0.118***	0.084***	0.092***
	(0.016)	(0.020)	(0.017)	(0.026)
Duty x medium corruption	0.006***	−0.009***	0.004**	0.007**
	(0.002)	(0.002)	(0.002)	(0.003)
Duty x high corruption	0.0002	−0.004	−0.0004	0.004
	(0.003)	(0.003)	(0.003)	(0.004)
Medium corruption	−0.006	0.077***	−0.001	0.076**
	(0.018)	(0.023)	(0.019)	(0.036)
High corruption	0.005	0.050	0.021	0.203***
	(0.026)	(0.032)	(0.029)	(0.046)
Country Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
p-value: duty*lowDem = duty*medDem	<0.01	<0.01	<0.01	<0.01
p-value: duty*medCorr = duty*highCorr	<0.01	0.0131	0.0517	0.049
Observations	612,761	397,532	314,121	565,589
R^2	0.038	0.030	0.063	0.044
Adjusted R^2	0.038	0.029	0.062	0.044

Notes: Data are for Swedish, Finnish, Norwegian, and Danish exports to all countries, from 2000 - 2021, aggregated at the HS4 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of versions of Specification 4 for the different export countries. The outcome variable is ln(Exports) - ln(Imports). Tertiles for the democracy and corruption variables are recalculated on a yearly basis. The variables are as described in the notes to Table 3. The control variables not shown in the table are the GDP and GDP per capita. The p-values shown in the bottom of the table refer to a two-sided t-test verifying the equality of the coefficients on different interaction terms. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

7 Conclusion

In this thesis we have provided an analysis of the possible determinants of customs duty evasion using mirror trade analysis, a prominent method in the trade literature. We build upon the existing body of literature which has consistently established that duty rate increases lead to an increase in tariff evasion and we provide novel results which suggest that this effect is much stronger for countries with lower levels of democracy and higher levels of corruption. We find that while a one percentage point increase in the duty rate of a high democracy country can lead to an average 0.5% increase in undervaluation fraud, this effect can more than double for countries with weaker democracies. The results are robust to the addition of various controls and as well as two-way fixed effects. Furthermore, they hold not just for Swedish exports but across the Nordic region as well.

These findings are significant especially since developing countries, which tend score badly on democracy and corruption indices, also rely more heavily on import duties for tax revenue. It has been established that tariff rate increases can lead to worse macroeconomic outcomes like unemployment and decline in productivity (Furceri et al., 2018). Our results further illustrate how engaging in this form of protectionism can be even more detrimental to the economy by decreasing revenue to the government due to evasion. As such, this thesis is also policy-relevant and provides support for moving towards more trade liberalisation. Instead of increasing tariffs to increase revenue to the government, a more effective solution would be to improve enforcement and monitoring in order to collect the duties which are already in place. Import countries can also set up monitoring systems which employ the mirror trade statistics methodology used in this paper for the purpose of identifying and preventing possible instances of fraud.

Lastly, this paper opens up a number of avenues for further research, perhaps most interesting of which is an investigation into how stark changes in governance, such as coups, can affect customs evasion. We provide some initial exploratory analysis for this in Appendix H. Additionally, conducting a case study analysis using more granular data, either on a monthly or even transactional level could provide further insights into the mechanisms which drive our overall results.

References

Afontsev, S. (2012). Political Economy of Tariff Unification The case of Russia. Economics Education and Research Consortium.

Alfieri, A., & Cirera, X. (2007). Fiscal Policy and Tax Incidence: Mozambique and Regional Integration. National Directorate of Studies, Policy Analysis Ministry of Planning, and Development, Republic of Mozambique. www.mpd.gov.mz

Berger, H., & Nitsch, V. (2008). Gotcha!: A Profile of Smuggling in International Trade. Center for Economic Studies and ifo Institute (CESifo). http://hdl.handle.net/10419/26520

Bhagwati, J. (1964). On the Underinvoicing of Imports. Bulletin of the Oxford University Institute of Economics & Statistics. *27*, 389–397. https://doi.org/10.1111/j.1468-0084.1964.mp27004007.x

Biersteker, T. J. (1990). Reducing the Role of the State in the Economy: A Conceptual Exploration of IMF and World Bank Prescriptions. International Studies Quarterly. *34*, 477. https://doi.org/10.2307/2600608

Bussy, A. (2020). Corporate Tax Evasion: Evidence from International Trade. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3607016

Bussy, A. (2021). Tariff Evasion with Endogenous Enforcement. Economics Letters. *207*, 110046. https://doi.org/10.1016/j.econlet.2021.110046

Cantens, T., Ireland, R., Raballand, G., & Bank, W. (2012). Reform by Numbers: Measurement Applied to Customs and Tax Administrations in Developing Countries. World Bank Publications. https://books.google.fr/books?id=RdEwI5W1mWcC

Cariolle, J., Chalendard, C., Geourjon, A., & Laporte, B. (2019). Measuring and Improving the Performance of Customs Valuation Controls: An illustration with Gabon. The World Economy. *42*, 1850–1872. https://doi.org/10.1111/twec.12759

Carrère, C., & Grigoriou, C. (2015). Can Mirror Data Help to Capture Informal International Trade? Fondation pour les études et recherches sur le développement international (FERDI). https://ferdi.fr/dl/df-6iH6FxjdWS8K1vAs43xfqnwQ/ferdi-p123-can-mirror-data-help-to-capture-informal-international-trade.pdf

Chalendard, C., Raballand, G., & Rakotoarisoa, A. (2019). The Use of Detailed Statistical Data in Customs Reforms: The Case of Madagascar. Development Policy Review. *37*, 546–563. https://doi.org/10.1111/dpr.12352

Chen, C., Jiang, Z., Li, N., Wang, H., Wang, P., Zhang, Z., Zhang, C., Ma, F., Huang, Y., Lu, X., Wei, J., Qi, J., & Chen, W.-Q. (2022). Advancing UN Comtrade for Physical Trade Flow Analysis: Review of Data Quality Issues and Solutions. Resources, Conservation and Recycling. *186*, 106526. https://doi.org/10.1016/j.resconrec.2022.106526

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Alizada, N., Altman, D., Bernhard, M., Cornell, A., Fish, M. S., Gastaldi, L., Gjerløw, H., Glynn, A., Hicken, A., Hindle, G., Ilchenko, N., Krusell, J., Luhrmann, A., Maerz, S. F., … Ziblatt., D.

(2022a). V-Dem Country-Year/Country-Date Dataset v12. https://www.v-dem.net/en/data/data-version-12/

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Altman, D., Bernhard, M., Cornell, A., Fish, M. S., Gastaldi, L., Gjerløw, H., Glynn, A., Hicken, A., Lührmann, A., Maerz, S. F., Marquardt, K. L., McMann, K., Mechkova, V., Paxton, P., ... Ziblatt, D. (2022b). V-Dem Codebook v12. https://www.v-dem.net/en/data/data-version-12/

Dahle, A. B., Thomasen, J., & Østereng, H. K. (1998). The Mirror Statistics Exercise between the Nordic Countries - 1995. Statistics Norway. https://www.ssb.no/a/histstat/doc/doc_199807.pdf

European Commission. (1987). Council Regulation on the Tariff and Statistical Nomenclature and on the Common Customs Tariff. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:31987R2658&from=en#d1e32-7-1

European Commission. (2020). Customs. European Commission, Directorate-General for Communication. https://doi.org/10.2775/851091

European Commission. (2021). Export Procedure. https://taxation-customs.ec.europa.eu/customs-4/customs-procedures-import-and-export-0/what-exportation/export-procedure_enExportprocedure

European Commission. (2022). Compendium of Customs Valuation Texts. https://taxation-customs.ec.europa.eu/system/files/2022-07/2022\%20EU\%20Valuation\%20Compendium\%20EN.pdf

European Commission. (2023). Monitoring Container Traffic and Analysing Risk. https://joint-research-centre.ec.europa.eu/scientific-activities-z/monitoring-container-traffic-and-analysing-risk_en

Fisman, R., & Wei, S. (2004). Tax Rates and Tax Evasion: Evidence from "Missing Imports" in China. Journal of Political Economy. *112*, 471–496. https://doi.org/10.1086/381476

Furceri, D., Hannan, S. A., Ostry, J. D., & Rose, A. K. (2018). Macroeconomic Consequences of Tariffs. (25402). https://doi.org/10.3386/w25402

Goemans, H. E., Gleditsch, K. S., & Chiozza, G. (2009). Introducing Archigos: A Dataset of Political Leaders. Journal of Peace Research. *46*(2), 269–283. Retrieved March 1, 2023, from http://www.jstor.org/stable/25654384

Hamanaka, S. (2011). Utilizing the Multiple Mirror Technique to Assess the Quality of Cambodian Trade Statistics. Asian Development Bank.

Hummels, D., & Lugovskyy, V. (2006). Are Matched Partner Trade Statistics a Usable Measure of Transportation Costs?*. Review of International Economics. *14*, 69–86. https://doi.org/10.1111/j.1467-9396.2006.00561.x

Hörnqvist, M. (2006). Risk Governance of the Swedish Customs Service: Negotiating Security, Efficiency, and Poor Information. Journal of Scandinavian Studies in Criminology and Crime Prevention. *7*, 23–44. https://doi.org/10.1080/14043850500522725

Javorcik, B. S., & Narciso, G. (2008). Differentiated Products and Evasion of Import Tariffs. Journal of International Economics. *76*, 208–222. https://doi.org/10.1016/j.jinteco.2008.07.002

Javorcik, B. S., & Narciso, G. (2017). WTO Accession and Tariff Evasion. Journal of Development Economics. *125*, 59–71. https://doi.org/10.1016/j.jdeveco.2016.11.001

Jean, S, & Mitaritonna, C. (2010). Determinants and Pervasiveness of the Evasion of Customs Duties. AGFOODTRADE (New Issues in Agricultural, Food and Bioenergy Trade). https://ageconsearch.umn.edu/record/115432/files/AgFoodTradeWP2010-05_Jean_Mitaritonna.pdf

Levin, J., & Widell, L. M. (2014). Tax Evasion in Kenya and Tanzania: Evidence from Missing Imports. Economic Modelling. *39*, 151–162. https://doi.org/10.1016/j.econmod.2014.02.021

Lyons, T. (2018). EU Customs Law. Oxford University Press.

Markhonko, V. (2014). Asymmetries in Official International Trade Statistics and Analysis of Globalization. Aguascalientes, Mexico.

Martinez-Vazquez, J., Vulovic, V., & Liu, Y. Direct Versus Indirect Taxation: Trends, theory, and economic significance. In: The Elgar Guide to Tax Systems (E. Albi & J. Martinez-Vazquez, Eds.). Ed. by Albi, E., & Martinez-Vazquez, J. Edward Elgar Publishing, 2011, pp. 37–92.

Nitsch, V. (2015). Trillion Dollar Estimate: Illicit Financial Flows from Developing Countries. Darmstadt. http://tuprints.ulb.tu-darmstadt.de/5437/

Nitsch, V. (2017). Trade Misinvoicing in Developing Countries. Centre for Global Development. https://www.cgdev.org/sites/default/files/trade-misinvoicing-developing-countries.pdf

OECD. (2023). Tax revenue (indicator). https://data.oecd.org/tax/tax-revenue.htm

Perotta, D, Cerasa, A, Torti, F, & Riani, M. (2020). The Robust Estimation of Monthly Prices of Goods Traded by the European Union: A Technical Guide. European Commission Joint Research Council. https://doi.org/10.2760/635844

Rijkers, B., Baghdadi, L., & Raballand, G. (2017). Political Connections and Tariff Evasion Evidence from Tunisia. The World Bank Economic Review. *31*, 459–482. https://doi.org/10.1093/wber/lhv061

Statistics Sweden. (2018). Quality Declaration: Foreign Trade - Exports and Imports of Goods. Statistikmyndigheten. https://www.scb.se/contentassets/4584f3f1df19400a885428aaa914d06e/ha0201_kd_2018_eng_ma_180222.pdf

Tax Justice Network. (2021). The State of Tax Justice 2021. https://taxjustice.net/wp-content/uploads/2021/11/State_of_Tax_Justice_Report_2021_ENGLISH.pdf

United Nations. (1997). UN Comtrade. Retrieved March 1, 2023, from https://comtrade.un.org/

Vargas, M. (2022). Tradestatistics: Open Trade Statistics API Wrapper and Utility Program.

Wahyudi, I. T., & Firdiansyah, A. (2019). A Perspective on Imposing Import Duty on Digital Goods in Indonesia. Customs Research and Applications Journal-CRAJ. *64*. https://craj.pknstan.ac.id/crajjournal/article/download/21/3

World Bank. (2001). World Integrated Trade Solutions (WITS). Retrieved March 16, 2023, from https://wits.worldbank.org/Default.aspx?lang=en

World Bank. (2023). Taxes on International Trade (% of Revenue) - High Income, Low & Middle Income, Sub-Saharan Africa. https://data.worldbank.org/indicator/GC.TAX.INTT.RV.ZS?end=2020&locations=XD-XO-ZG&most_recent_value_desc=false&start=2000

World Trade Organisation. (1994). General Agreement on Tariffs and Trade 1994. Article VII: Valuation for Customs Purposes. https://www.wto.org/english/docs_e/legal_e/gatt47_01_e.htm#articleVII

World Trade Organisation. (2000). WTO Tariff Download Facility. http://tariffdata.wto.org/

World Trade Organisation. (2009). WTO Regional Trade Agreements. http://rtais.wto.org/UI/PublicMaintainRTAHome.aspx

World Trade Organisation. (2023). HS Tracker. https://hstracker.wto.org/?_inputs_&page=\%22visualizer\%22&sidebarCollapsed=false&hscombined_dt_table_row_last_clicked=587&hscode=\%22060310\%22

Yu, D. (2008). The Harmonized System - Amendments and Their Impact on WTO Members' Schedules. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.1117865

A Additional figures

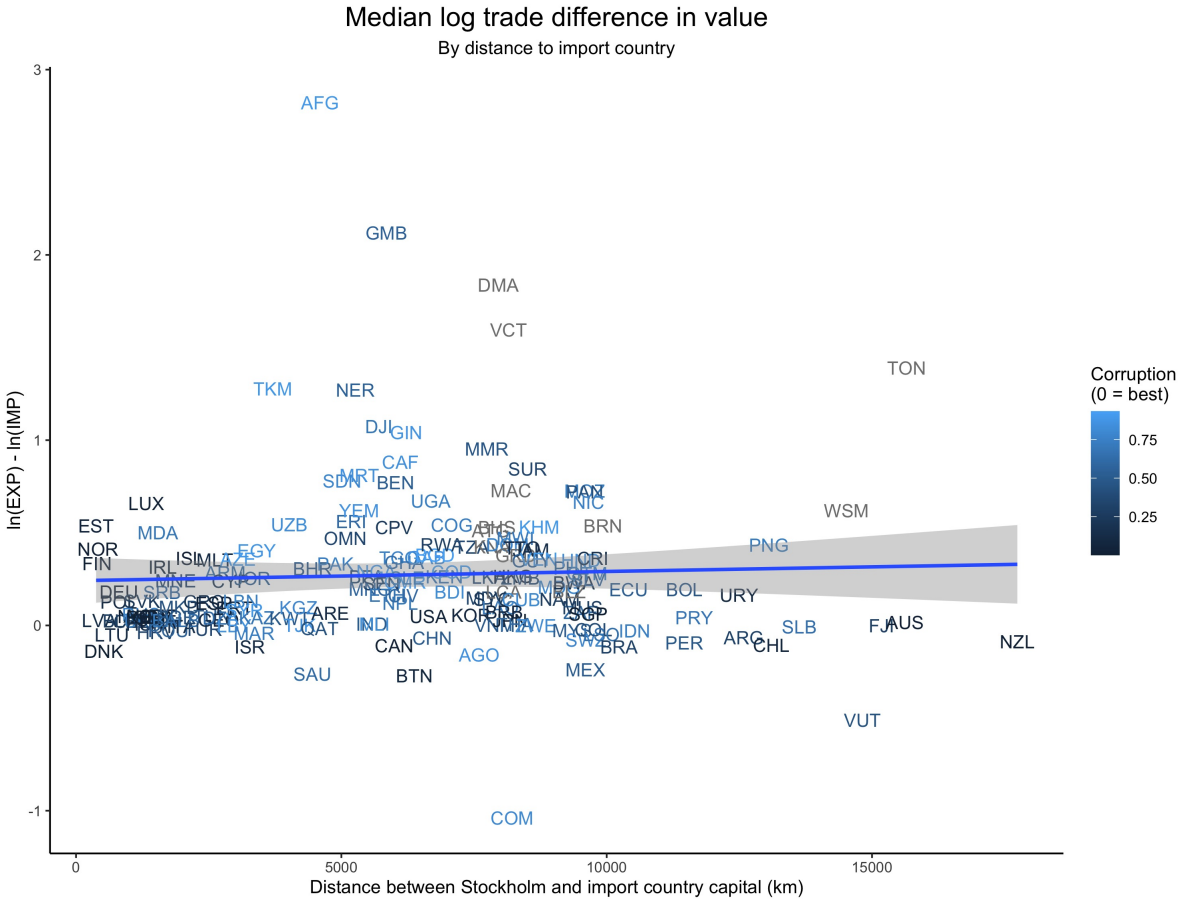

Figure A1: Median trade gap by distance from Stockholm, 2000 - 2021

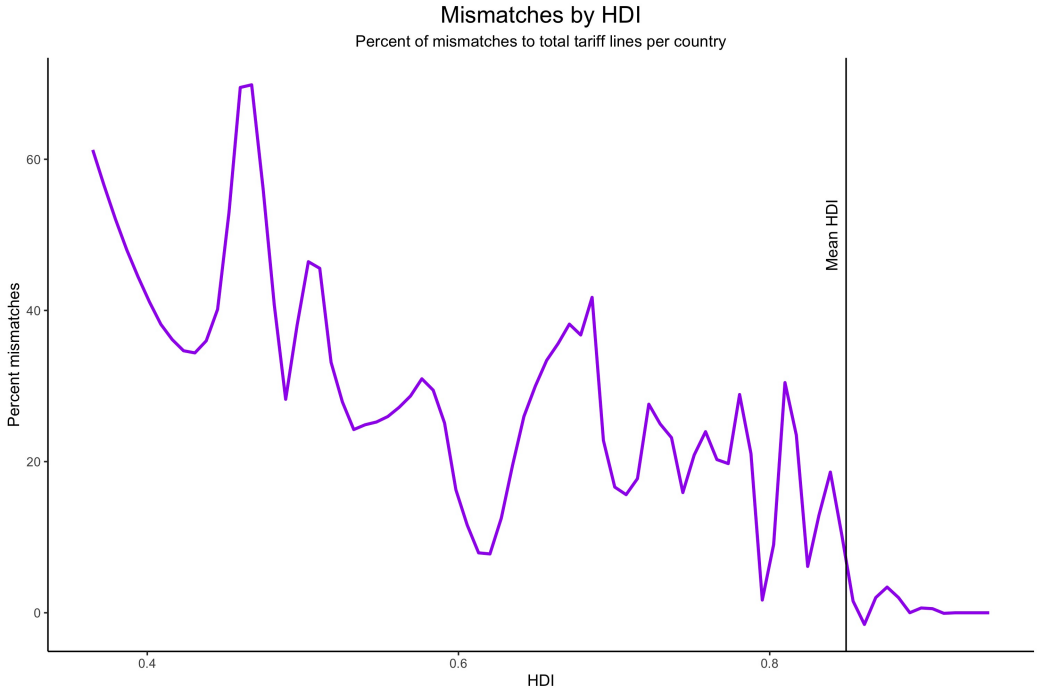

Figure A2: Percent of tariff lines with HS mismatches by country HDI

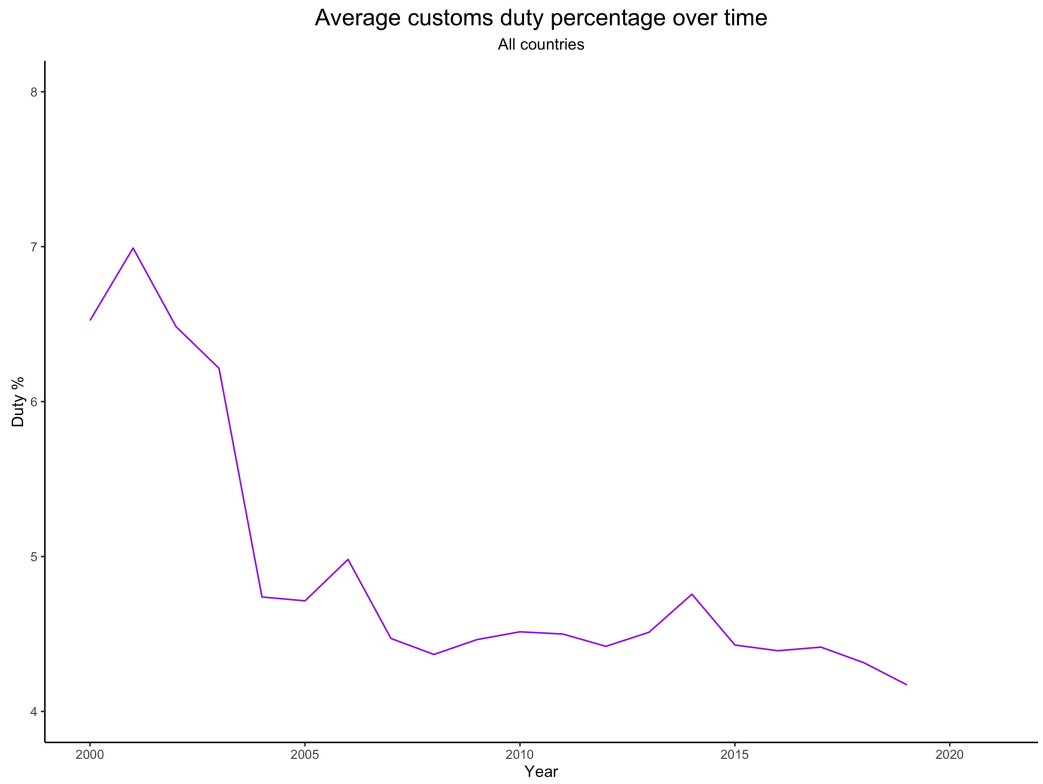

Figure A3: Trend in customs duty rates charged on all Swedish exports, 2000 - 2021

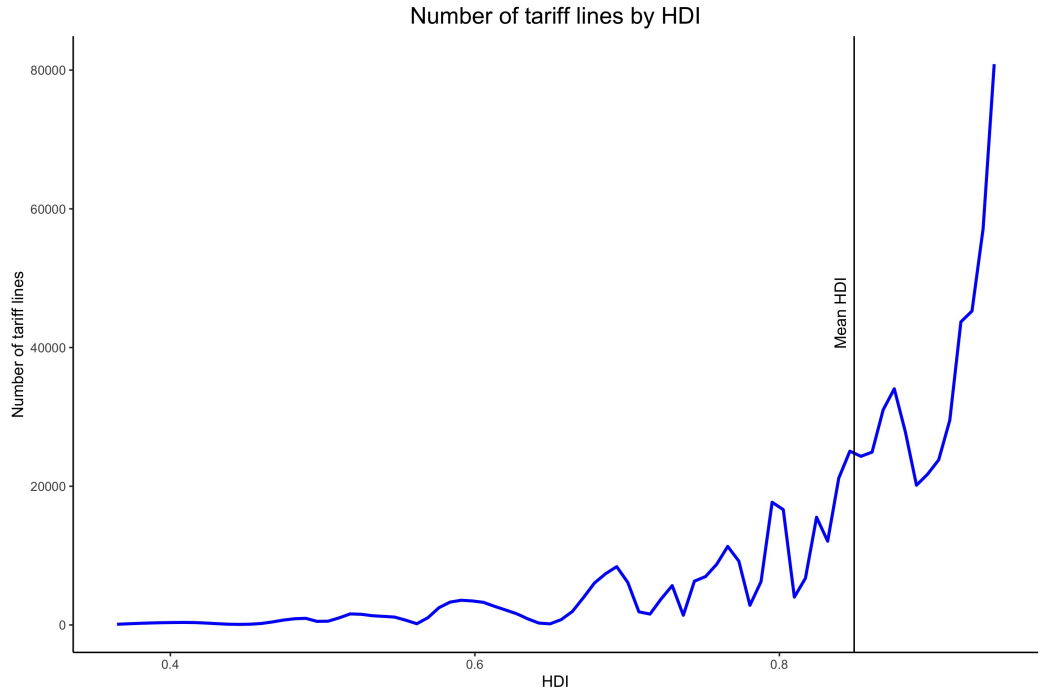

Figure A4: Number of HS6 tariff lines per country ordered by HDI

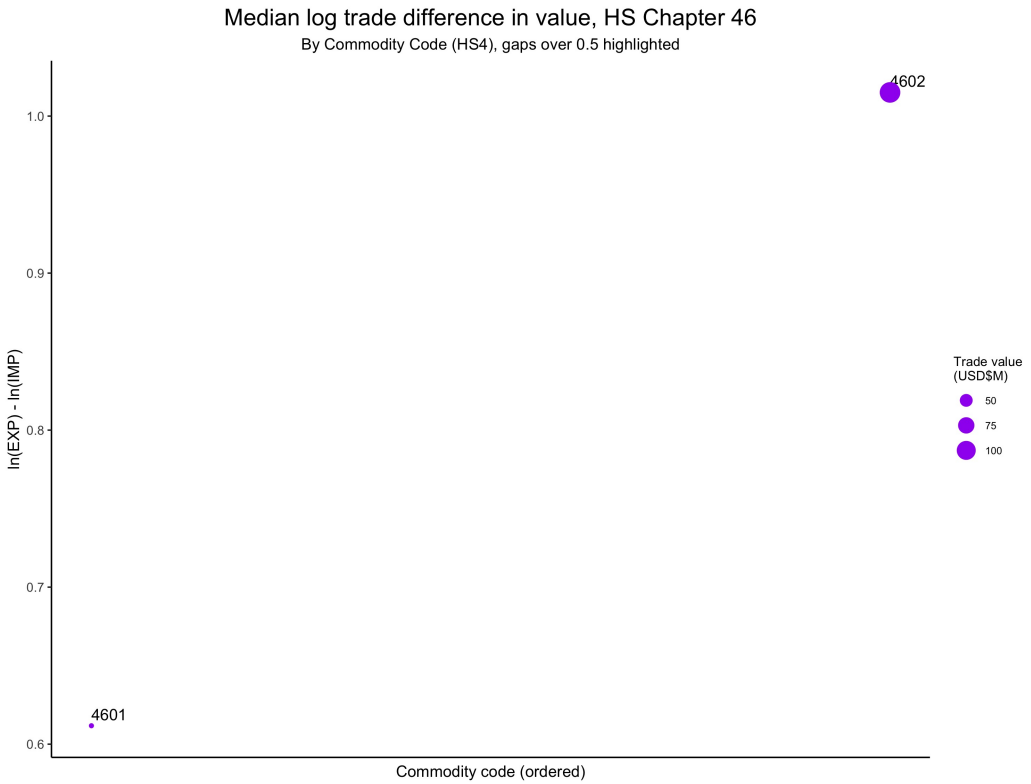

Figure A5: Median trade gap within HS Chapter 46, 2000 - 2021

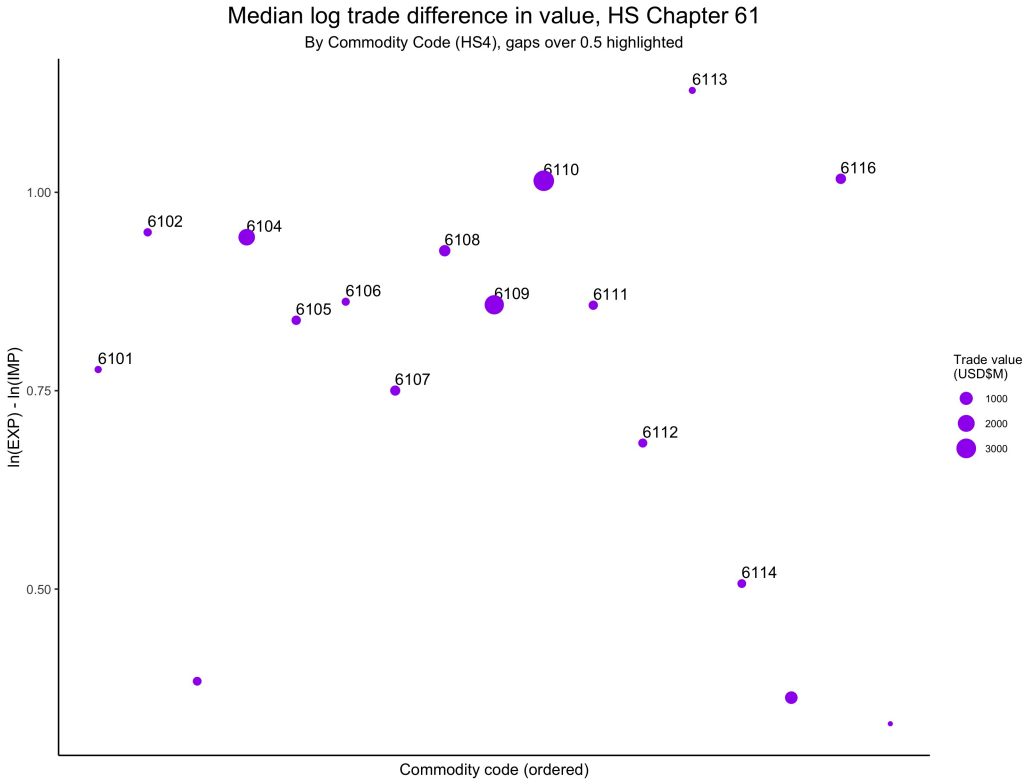

Figure A6: Median trade gap within HS Chapter 61, 2000 - 2021

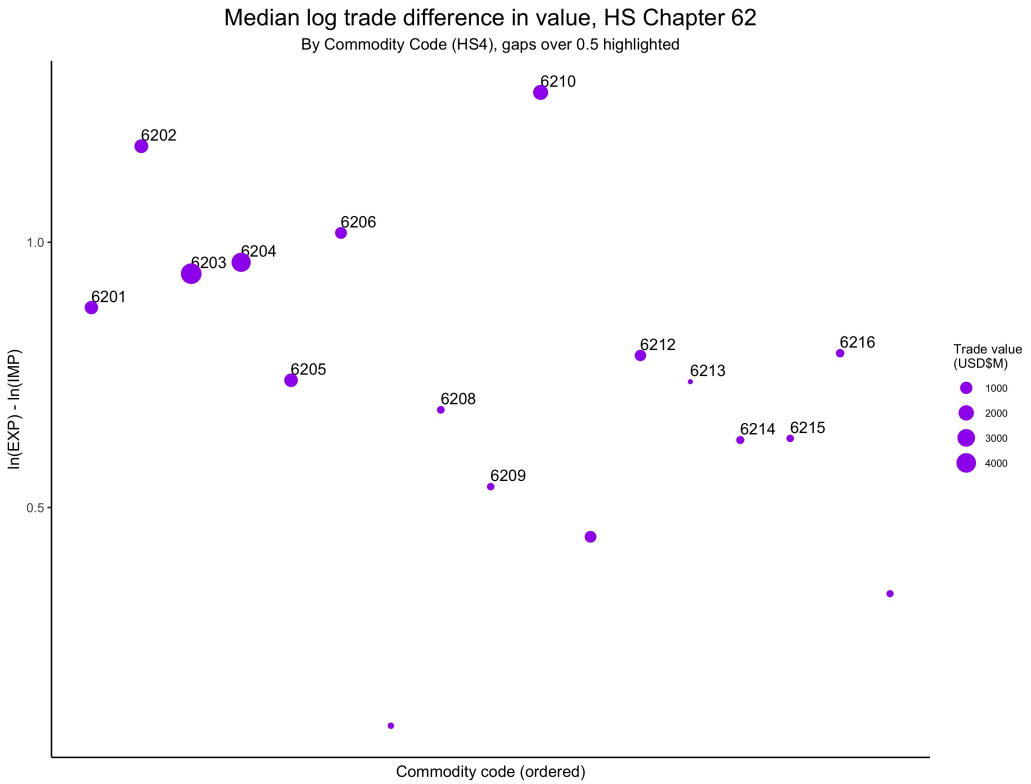

Figure A7: Median trade gap within HS Chapter 62, 2000 - 2021

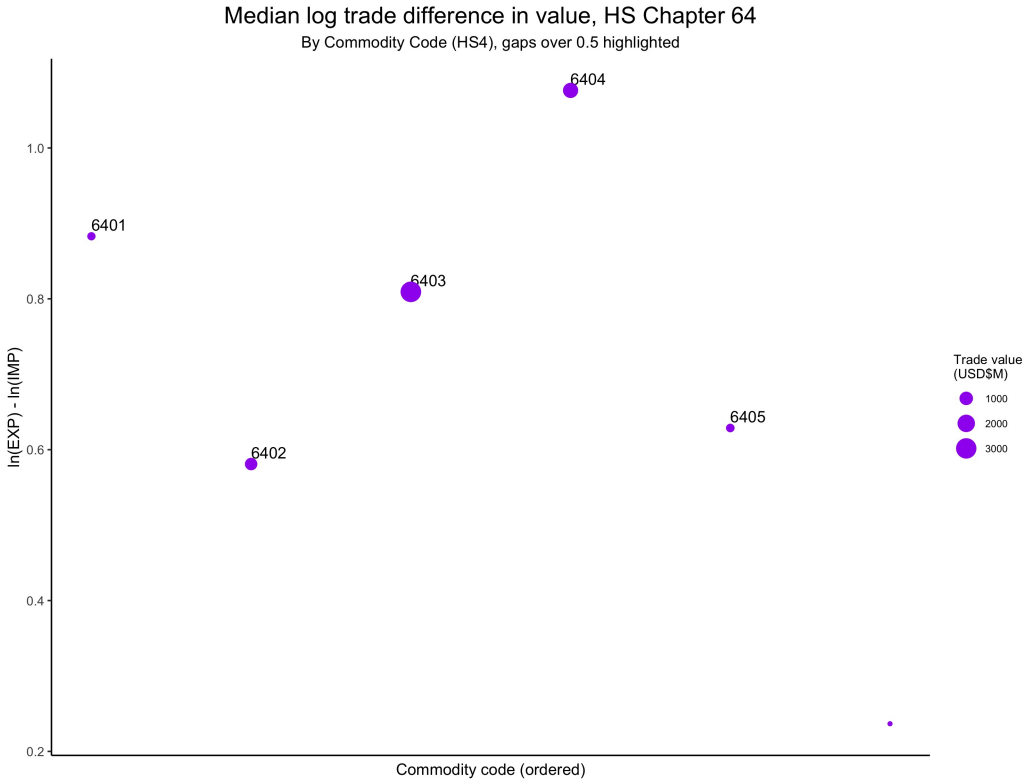

Figure A8: Median trade gap within HS Chapter 64, 2000 - 2021

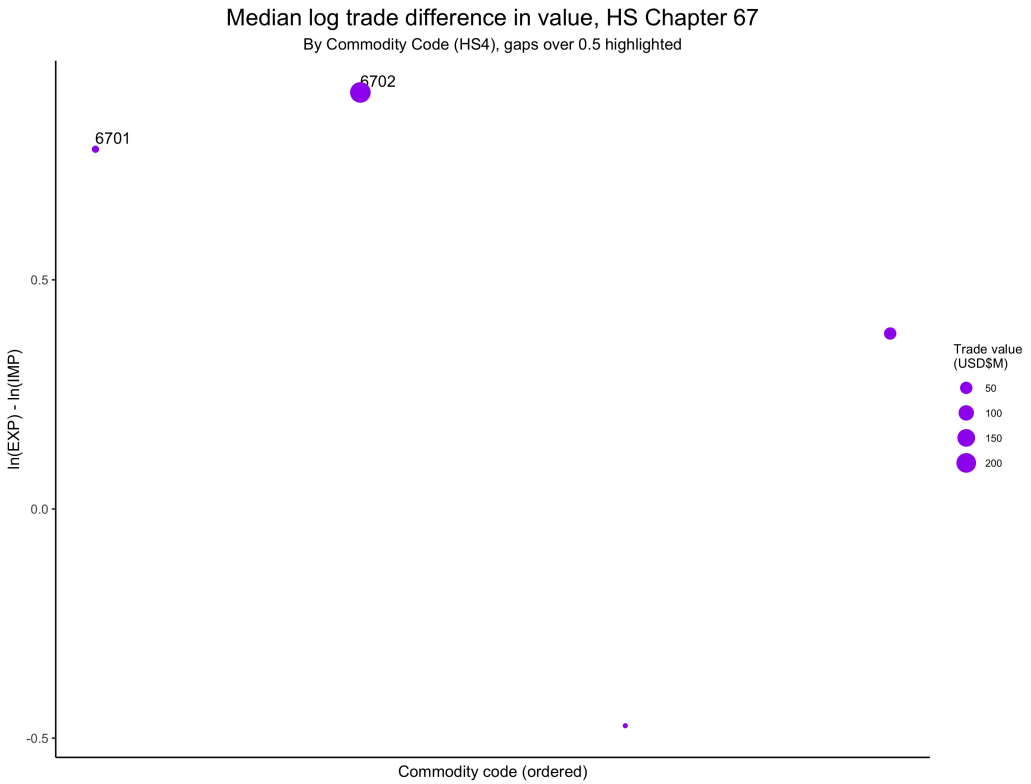

Figure A9: Median trade gap within HS Chapter 67, 2000 - 2021

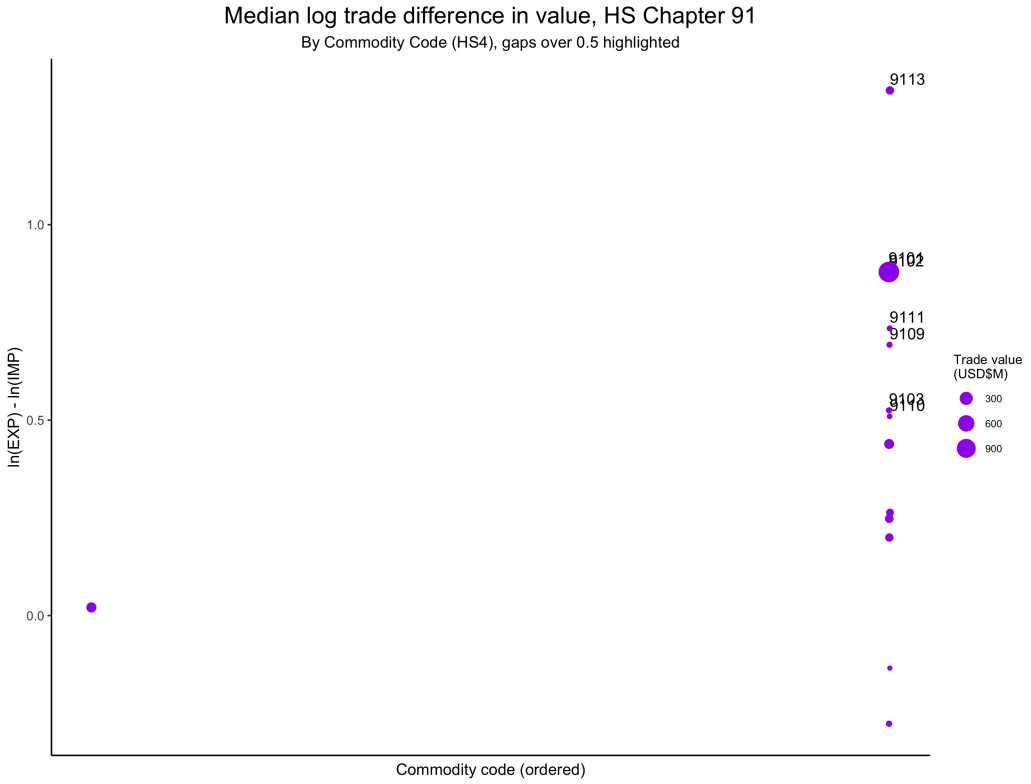

Figure A10: Median trade gap within HS Chapter 91, 2000 - 2021

B Tertile balance

Import Country	Democracy Tertile Changes
Argentina	2
Armenia	1
Bangladesh	1
Bolivia (Plurinational State of)	1
Botswana	2
Brazil	1
Côte d'Ivoire	5
China, Hong Kong SAR	3
Croatia	4
Dominican Rep.	1
Ecuador	2
Fiji	2
Gambia	1
Georgia	1
Ghana	5
Honduras	1
Hungary	1
Jamaica	1
Kenya	3
Kuwait	4
Kyrgyzstan	1
Madagascar	1
Malaysia	1
Maldives	3
Mali	2
Malta	3
Mauritius	1
Mozambique	1
Nepal	3
Nicaragua	1
Niger	2
Peru	4
Philippines	1
Poland	1
Rep. of Korea	2
Romania	1
Singapore	4
Slovenia	1
South Africa	3
Sri Lanka	2
Thailand	3
Trinidad and Tobago	4
Tunisia	2
Turkey	1
Ukraine	2
Venezuela	1

Import Country	Corruption Tertile Changes
Albania	1
Armenia	1
Benin	3
Botswana	1
Brazil	2
Bulgaria	1
Burkina Faso	2
Burundi	1
Côte d'Ivoire	2
Cabo Verde	2
Croatia	2
Cyprus	1
El Salvador	9
Fiji	1
Gambia	2
Georgia	1
Ghana	1
Greece	2
Hungary	3
India	1
Kenya	1
Lesotho	1
Malawi	2
Malaysia	2
Maldives	2
Mauritius	3
Mexico	2
Mongolia	2
Myanmar	2
Niger	3
Paraguay	2
Philippines	4
Rep. of Moldova	2
Romania	1
Saudi Arabia	2
Senegal	2
Slovakia	6
Slovenia	1
Solomon Islands	1
South Africa	5
Thailand	3
Tunisia	3
Turkey	2
Ukraine	2
United Rep. of Tanzania	1
Zambia	4

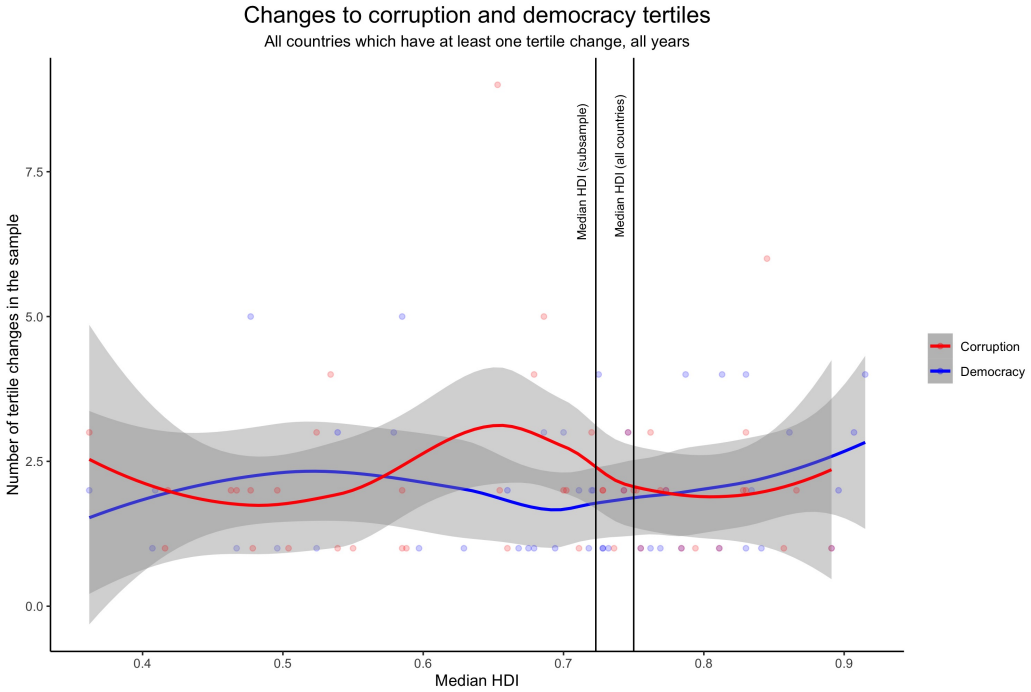

Figure B1: Median trade gap within HS Chapter 91, 2000 - 2021

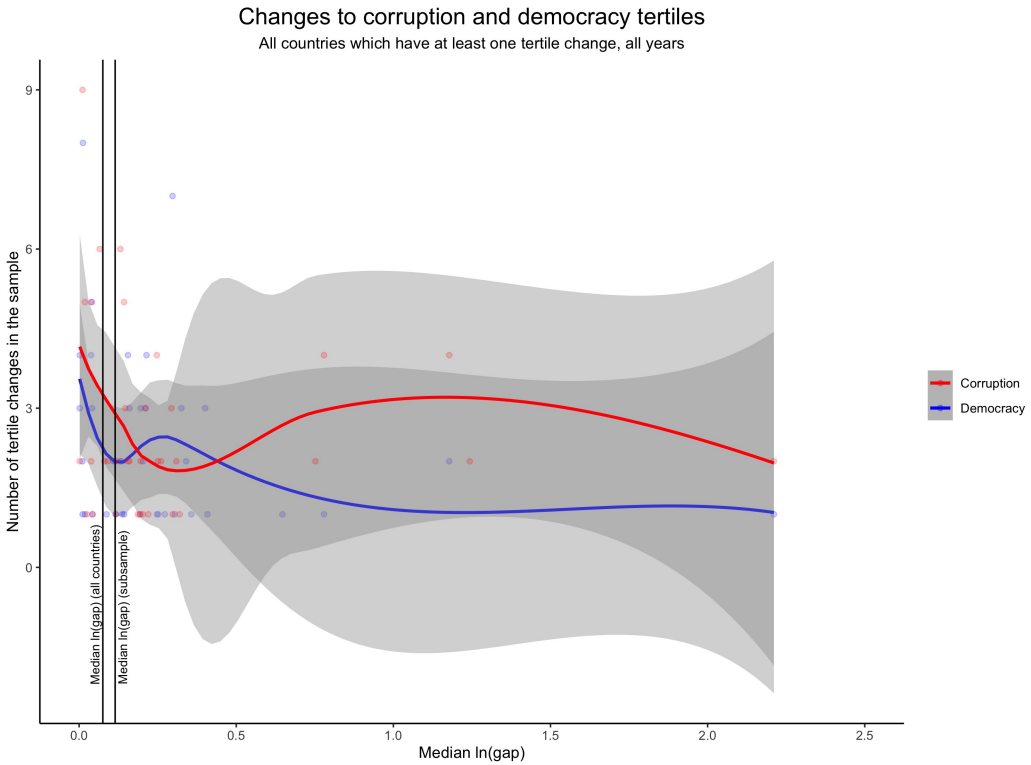

Figure B2: Median trade gap within HS Chapter 91, 2000 - 2021

C Results at different HS granularity levels

C.1 Comparison tables

Table C1: Sweden - Baseline regressions at different HS levels

	Dependent variable:		
	lnGap_USD		
	(HS6)	(HS4)	(HS2)
Duty Percentage	0.006***	0.007***	0.008***
	(0.0004)	(0.001)	(0.001)
Democracy (1 = best)	−0.085***	−0.090***	−0.108***
	(0.012)	(0.014)	(0.027)
Corruption (0 = best)	0.202***	0.234***	0.173***
	(0.013)	(0.017)	(0.033)
Distance to capital (km)	−0.00001***	−0.00001***	−0.00001***
	(0.00000)	(0.00000)	(0.00000)
GDP	−0.000***	−0.000***	−0.000***
	(0.000)	(0.000)	(0.000)
GDP per capita	0.00000***	0.00000***	0.00000***
	(0.00000)	(0.00000)	(0.00000)
Constant	0.199***	0.116***	0.047*
	(0.011)	(0.014)	(0.027)
Observations	1,504,167	612,761	93,399
R^2	0.002	0.003	0.004
Adjusted R^2	0.002	0.003	0.004

Notes: Data are for Swedish exports to all countries, from 2000 - 2021. Robust standard errors reported in parentheses. This table reports the results of Specification 2, with the data aggregated at the HS2, HS4, and HS6 levels of the harmonised system. The outcome variable is ln(Exports) - ln(Imports). The duty percentage refers to the amount of customs duty levied on imports of Swedish goods into a given import country. As explained in Section 4.2.2, we assume that any applicable Free Trade Agreements are fully utilised. The democracy variable, from the VDEM dataset, is a composite of two indices, one which measures electoral democracy through scores for freedom of association, freedom of expression, the quality of the electoral process, the pervasiveness of the electoral system in selecting a country's leaders, and the extent to which suffrage is universal, and the other which includes measures of judicial and legislative constraints on executive power, as well as useful measures for state capacity. The corruption variable is an aggregate of measures for the average level of bribery in the public sector as well as measures for theft of state property by public sector officials. Both the democracy and corruption variables are from the VDEM dataset. The distance variable records the distance between Stockholm and the import country's capital. The GDP and GDP per capita relate to the import country. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

Table C2: Sweden - Fixed Effects specifications at different HS levels

	Dependent variable:		
	lnGap_USD		
	(HS6)	(HS4)	(HS2)
Duty Percentage	0.009***	0.009***	0.007***
	(0.001)	(0.001)	(0.001)
Democracy (1 = best)	−0.211***	−0.100*	−0.046
	(0.038)	(0.052)	(0.098)
Corruption (0 = best)	0.168***	0.232***	0.278***
	(0.046)	(0.060)	(0.107)
Country Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Observations	1,504,167	612,761	93,399
R^2	0.035	0.038	0.049
Adjusted R^2	0.035	0.038	0.048

Notes: Data are for Swedish exports to all countries, from 2000 - 2021, with the data aggregated at the HS2, HS4, and HS6 levels of the harmonised system. Robust standard errors reported in parentheses. Robust standard errors reported in parentheses. This table reports the results of Specification 3, with all the covariates included along with both Country FE and Year FE. The outcome variable is ln(Exports) - ln(Imports). The variables are as described in the notes to Table 3. The control variables not shown in the table are the GDP and GDP per capita. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

Table C3: Regressions with interaction terms at different HS levels

	Dependent variable:		
	lnGap_USD		
	(HS6)	(HS4)	(HS2)
Duty Percentage	0.013***	0.005***	0.003**
	(0.001)	(0.001)	(0.002)
Duty x low democracy	−0.001	0.005**	0.010**
	(0.002)	(0.002)	(0.004)
Duty x medium democracy	−0.005***	−0.001	0.008**
	(0.001)	(0.002)	(0.003)
Low democracy	0.147***	0.072***	−0.001
	(0.018)	(0.024)	(0.044)
Medium democracy	0.168***	0.084***	0.008
	(0.012)	(0.016)	(0.030)
Duty x medium corruption	−0.001	0.006***	−0.001
	(0.002)	(0.002)	(0.003)
Duty x high corruption	−0.009***	0.0002	−0.003
	(0.002)	(0.003)	(0.004)
Medium corruption	−0.010	−0.006	0.013
	(0.014)	(0.018)	(0.036)
High corruption	−0.005	0.005	0.010
	(0.020)	(0.026)	(0.049)
Country Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Controls	Yes	Yes	Yes
p-value: duty*lowDem = duty*medDem	<0.01	<0.01	0.2205
p-value: duty*medCorr = duty*highCorr	<0.01	<0.01	0.516
Observations	1,504,167	612,761	93,399
R^2	0.035	0.038	0.050
Adjusted R^2	0.035	0.038	0.048

Notes: Data are for Swedish exports to all countries, from 2000 - 2021, with the data aggregated at the HS2, HS4, and HS6 levels of the harmonised system. Robust standard errors reported in parentheses. This table reports the results of Specification 4 with all interaction terms included. The outcome variable is ln(Exports) - ln(Imports). Tertiles for the democracy and corruption variables are recalculated on a yearly basis. The variables are as described in the notes to Table 3. The control variables not shown in the table are the GDP and GDP per capita. The p-values shown in the bottom of the table refer to a two-sided t-test verifying the equality of the coefficients on different interaction terms. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

C.2 HS6 - Full results

Table C4: Sweden - Regressions with different covariates

	Dependent variable:					
	lnGap_USD					
	(1)	(2)	(3)	(4)	(5)	(6)
Duty Percentage	0.001***	0.001	0.0004	0.0004	0.004***	0.006***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0004)	(0.0004)
Democracy (1 = best)		−0.023***		−0.007	−0.077***	−0.085***
		(0.008)		(0.011)	(0.012)	(0.012)
Corruption (0 = best)			0.029***	0.023**	0.029***	0.202***
			(0.008)	(0.011)	(0.011)	(0.013)
Distance to capital (km)					−0.00002***	−0.00001***
					(0.00000)	(0.00000)
GDP						−0.000***
						(0.000)
GDP per capita						0.00000***
						(0.00000)
Constant	0.217***	0.233***	0.212***	0.218***	0.322***	0.199***
	(0.002)	(0.006)	(0.002)	(0.010)	(0.010)	(0.011)
Observations	1,505,209	1,504,167	1,504,429	1,504,167	1,504,167	1,504,167
R^2	0.00001	0.00001	0.00002	0.00002	0.001	0.002
Adjusted R^2	0.00001	0.00001	0.00001	0.00001	0.001	0.002

Notes: Data are for Swedish exports to all countries, from 2000 - 2021, aggregated at the HS6 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of versions of Specification 2. The outcome variable is ln(Exports) - ln(Imports). The duty percentage refers to the amount of customs duty levied on imports of Swedish goods into a given import country. As explained in Section 4.2.2, we assume that any applicable Free Trade Agreements are fully utilised. The democracy variable, from the VDEM dataset, is a composite of two indices, one which measures electoral democracy through scores for freedom of association, freedom of expression, the quality of the electoral process, the pervasiveness of the electoral system in selecting a country's leaders, and the extent to which suffrage is universal, and the other which includes measures of judicial and legislative constraints on executive power, as well as useful measures for state capacity. The corruption variable is an aggregate of measures for the average level of bribery in the public sector as well as measures for theft of state property by public sector officials. Both the democracy and corruption variables are from the VDEM dataset. The distance variable records the distance between Stockholm and the import country's capital. The GDP and GDP per capita relate to the import country. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

Table C5: Sweden - Fixed Effects specifications

	Dependent variable:			
	lnGap_USD			
	(6)	(7)	(8)	(9)
Duty Percentage	0.006***	0.008***	0.006***	0.009***
	(0.0004)	(0.001)	(0.0004)	(0.001)
Democracy (1 = best)	−0.085***	−0.445***	−0.054***	−0.211***
	(0.012)	(0.036)	(0.012)	(0.038)
Corruption (0 = best)	0.202***	0.030	0.199***	0.168***
	(0.013)	(0.045)	(0.013)	(0.046)
Distance to capital (km)	−0.00001***		−0.00001***	
	(0.00000)		(0.00000)	
Constant	0.199***			
	(0.011)			
Country Fixed Effects	No	Yes	No	Yes
Year Fixed Effects	No	No	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	1,504,167	1,504,167	1,504,167	1,504,167
R^2	0.002	0.035	0.015	0.035
Adjusted R^2	0.002	0.035	0.015	0.035

Notes: Data are for Swedish exports to all countries, from 2000 - 2021, aggregated at the HS6 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of Specifications 2 and 3. The outcome variable is ln(Exports) - ln(Imports). Column (6) presents our preferred version of Specification 2, i.e. with all the covariates included, but with no Fixed Effects. Columns (7) to (9) add Country FE, Year FE, and then both. The variables are as described in the notes to Table 3. The control variables not shown in the table are the GDP and GDP per capita. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

Table C6: Sweden - Regressions with interaction terms

	Dependent variable:		
	lnGap_USD		
	(10)	(11)	(12)
Duty Percentage	0.012***	0.012***	0.012***
	(0.001)	(0.001)	(0.001)
Duty x low democracy	−0.005***		0.002
	(0.001)		(0.002)
Duty x medium democracy	−0.004***		−0.001
	(0.001)		(0.001)
Low democracy	0.068***		0.055***
	(0.019)		(0.019)
Medium democracy	0.123***		0.117***
	(0.010)		(0.010)
Duty x medium corruption		−0.003**	−0.004**
		(0.001)	(0.002)
Duty x high corruption		−0.010***	−0.012***
		(0.001)	(0.002)
Medium corruption		0.021	0.013
		(0.013)	(0.014)
High corruption		0.036*	0.027
		(0.019)	(0.019)
Country Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Controls	Yes	Yes	Yes
p-value: duty*lowDem = duty*medDem	0.5466		<0.01
p-value: duty*medCorr = duty*highCorr		<0.01	<0.01
Observations	1,504,167	1,504,429	1,504,167
R^2	0.035	0.035	0.035
Adjusted R^2	0.035	0.035	0.035

Notes: Data are for Swedish exports to all countries, from 2000 - 2021, aggregated at the HS6 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of versions of Specification 4 with different interaction terms included. The outcome variable is ln(Exports) - ln(Imports). Column (12), our preferred specification, includes interaction terms between the duty percentage and tertiles of both the democracy and corruption scores. Tertiles for the democracy and corruption variables are recalculated on a yearly basis. The variables are as described in the notes to Table 3. The control variables not shown in the table are the GDP and GDP per capita. The p-values shown in the bottom of the table refer to a two-sided t-test verifying the equality of the coefficients on different interaction terms. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

C.3 HS2 - Full results

Table C7: Sweden - Regressions with different covariates

	Dependent variable:					
	lnGap_USD					
	(1)	(2)	(3)	(4)	(5)	(6)
Duty Percentage	0.007***	0.006***	0.006***	0.006***	0.007***	0.008***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Democracy (1 = best)		−0.093***		−0.115***	−0.136***	−0.108***
		(0.019)		(0.027)	(0.027)	(0.027)
Corruption (0 = best)			0.053***	−0.032	−0.029	0.173***
			(0.020)	(0.029)	(0.029)	(0.033)
Distance to capital (km)					−0.00001***	−0.00001***
					(0.00000)	(0.00000)
GDP						−0.000***
						(0.000)
GDP per capita						0.00000***
						(0.00000)
Constant	0.088***	0.147***	0.075***	0.168***	0.209***	0.047*
	(0.005)	(0.014)	(0.006)	(0.023)	(0.023)	(0.027)
Observations	93,611	93,399	93,435	93,399	93,399	93,399
R^2	0.002	0.002	0.002	0.002	0.002	0.004
Adjusted R^2	0.002	0.002	0.002	0.002	0.002	0.004

Notes: Data are for Swedish exports to all countries, from 2000 - 2021, aggregated at the HS2 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of versions of Specification 2. The outcome variable is ln(Exports) - ln(Imports). The duty percentage refers to the amount of customs duty levied on imports of Swedish goods into a given import country. As explained in Section 4.2.2, we assume that any applicable Free Trade Agreements are fully utilised. The democracy variable, from the VDEM dataset, is a composite of two indices, one which measures electoral democracy through scores for freedom of association, freedom of expression, the quality of the electoral process, the pervasiveness of the electoral system in selecting a country's leaders, and the extent to which suffrage is universal, and the other which includes measures of judicial and legislative constraints on executive power, as well as useful measures for state capacity. The corruption variable is an aggregate of measures for the average level of bribery in the public sector as well as measures for theft of state property by public sector officials. Both the democracy and corruption variables are from the VDEM dataset. The distance variable records the distance between Stockholm and the import country's capital. The GDP and GDP per capita relate to the import country. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

Table C8: Sweden - Fixed Effects specifications

	Dependent variable:			
	lnGap_USD			
	(6)	(7)	(8)	(9)
Duty Percentage	0.008***	0.007***	0.007***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
Democracy (1 = best)	−0.108***	−0.060	−0.112***	−0.046
	(0.027)	(0.092)	(0.027)	(0.098)
Corruption (0 = best)	0.173***	0.296***	0.230***	0.278***
	(0.033)	(0.105)	(0.034)	(0.107)
Distance to capital (km)	−0.00001***		−0.00001***	
	(0.00000)		(0.00000)	
Constant	0.047*			
	(0.027)			
Country Fixed Effects	No	Yes	No	Yes
Year Fixed Effects	No	No	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	93,399	93,399	93,399	93,399
R^2	0.004	0.049	0.012	0.049
Adjusted R^2	0.004	0.047	0.012	0.048

Notes: Data are for Swedish exports to all countries, from 2000 - 2021, aggregated at the HS2 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of Specifications 2 and 3. The outcome variable is ln(Exports) - ln(Imports). Column (6) presents our preferred version of Specification 2, i.e. with all the covariates included, but with no Fixed Effects. Columns (7) to (9) add Country FE, Year FE, and then both. The variables are as described in the notes to Table 3. The control variables not shown in the table are the GDP and GDP per capita. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

Table C9: Sweden - Regressions with interaction terms

	Dependent variable:		
	lnGap_USD		
	(10)	(11)	(12)
Duty Percentage	0.003**	0.004**	0.003**
	(0.002)	(0.002)	(0.002)
Duty x low democracy	0.009***		0.010**
	(0.003)		(0.004)
Duty x medium democracy	0.006***		0.008**
	(0.002)		(0.003)
Low democracy	0.004		−0.001
	(0.042)		(0.044)
Medium democracy	0.013		0.008
	(0.028)		(0.030)
Duty x medium corruption		0.005**	−0.001
		(0.002)	(0.003)
Duty x high corruption		0.006*	−0.003
		(0.003)	(0.004)
Medium corruption		0.004	0.013
		(0.034)	(0.036)
High corruption		−0.006	0.010
		(0.047)	(0.049)
Country Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Controls	Yes	Yes	Yes
p-value: duty*lowDem = duty*medDem	0.2735		0.2205
p-value: duty*medCorr = duty*highCorr		0.893	0.516
Observations	93,399	93,435	93,399
R^2	0.050	0.049	0.050
Adjusted R^2	0.048	0.048	0.048

Notes: Data are for Swedish exports to all countries, from 2000 - 2021, aggregated at the HS2 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of versions of Specification 4 with different interaction terms included. The outcome variable is ln(Exports) - ln(Imports). Column (12), our preferred specification, includes interaction terms between the duty percentage and tertiles of both the democracy and corruption scores. Tertiles for the democracy and corruption variables are recalculated on a yearly basis. The variables are as described in the notes to Table 3. The control variables not shown in the table are the GDP and GDP per capita. The p-values shown in the bottom of the table refer to a two-sided t-test verifying the equality of the coefficients on different interaction terms. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

D Results with different FTA assumptions

We run the below specification a number of times to test our assumptions about FTA usage.

$$y_{ijk} = \alpha_{ijk} + \beta Tax_{ijk} + X + \epsilon_{ijk} \tag{5}$$

Where X is a vector of control variables as specified in equation 1 and Tax is the average duty percentage for a given commodity code k, import country i, and year j.

As explained in Section 4.2.2, we change our assumptions about Tax and rerun the regression with the same covariates. This allows us to see the impact of assuming 0% FTA usage, 25%, 50%, 75%, and 100%. In the specifications used throughout this thesis, we assume 100% FTA usage.

Table D1: Sweden - Regressions with different FTA use assumptions

	Dependent variable:				
	lnGap_USD				
	(1)	(2)	(3)	(4)	(5)
Duty - No FTAs	0.004***				
	(0.001)				
Duty - 25 percent FTA usage		0.005***			
		(0.001)			
Duty - 50 percent FTA usage			0.007***		
			(0.001)		
Duty - 75 percent FTA usage				0.008***	
				(0.001)	
Duty - 100 percent FTA usage					0.007***
					(0.001)
Distance to capital (km)	−0.00001***	−0.00001***	−0.00001***	−0.00001***	−0.00001***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Democracy (1 = best)	−0.097***	−0.095***	−0.092***	−0.090***	−0.090***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Corruption (0 = best)	0.234***	0.229***	0.224***	0.224***	0.234***
	(0.019)	(0.018)	(0.018)	(0.017)	(0.017)
Constant	0.123***	0.121***	0.118***	0.115***	0.116***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	612,761	612,761	612,761	612,761	612,761
R^2	0.003	0.003	0.003	0.003	0.003
Adjusted R^2	0.003	0.003	0.003	0.003	0.003

Notes: Data are for Swedish exports to all countries, from 2000 - 2021, aggregated at the HS4 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of Specification 5. The outcome variable is ln(Exports) - ln(Imports). Column (5) presents our preferred version of specification, assuming 100% FTA usage. The variables are as described in the notes to Table 3. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

E Weighted OLS results

Table E1: Weighted and non-weighted OLS comparison

	Dependent variable:	
	lnGap_USD	
	(6) Unweighted	(6) Weighted
Duty Percentage	0.007***	0.009***
	(0.001)	(0.002)
Democracy (1 = best)	−0.090***	−0.210***
	(0.014)	(0.058)
Corruption (0 = best)	0.234***	0.242***
	(0.017)	(0.070)
Distance to capital (km)	−0.00001***	−0.00000
	(0.00000)	(0.00000)
GDP	−0.000***	−0.000***
	(0.000)	(0.000)
GDP per capita	0.00000***	−0.00000**
	(0.00000)	(0.00000)
Constant	0.116***	0.467***
	(0.014)	(0.057)
Observations	612,761	612,761
R^2	0.003	0.011
Adjusted R^2	0.003	0.011

Notes: Data are for Swedish exports to all countries, from 2000 - 2021, aggregated at the HS4 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of versions of Specification 2. The unweighted column presents the naive OLS results, whereas the weighted column gives each import country in the sample equal weight. The variables are as described in the notes to Table 3. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

F Swedish Free Trade Agreements (FTAs)

Import Country	First Year Validity
Andorra	Before 2000
Albania	2009
Algeria	2005
Armenia	2018
Austria	Before 2000
Bahamas	2009
Bosnia and Herzegovina	2015
Barbados	2009
Belgium	Before 2000
Bulgaria	2007
Botswana	2016
Belize	2009
Cameroon	2014
Canada	2017
Chile	2005
Colombia	2013
Comoros	2019
Costa Rica	2013
Côte D'ivoire	2016
Croatia	2013
Cyprus	2004
Czech Republic	2004
Germany	Before 2000
Denmark	Before 2000
Dominica	2009
Dominican Republic	2009
Ecuador	2013
El Salvador	2013
Estonia	2004
Egypt	2004
Faroe Islands	Before 2000
Fiji	2009
Finland	Before 2000
France	Before 2000
Grenada	2009
Georgia	2014
Ghana	2016
Greece	Before 2000

Import Country	First Year Validity
Guatemala	2013
Guyana	2009
Honduras	2013
Hungary	2004
Iceland	Before 2000
Ireland	Before 2000
Israel	2000
Italy	Before 2000
Jamaica	2009
Japan	2019
Jordan	2002
Korea, Republic of	2011
Poland	2004
Portugal	Before 2000
Romania	2007
Saint Kitts and Nevis	2009
Saint Lucia	2009
Saint Vincent and the Grenadines	2009
San Marino	2002
Serbia	2013
Solomon Islands	2020
Seychelles	2012
Singapore	2019
Slovakia	2004
Spain	Before 2000
Suriname	2009
Switzerland	Before 2000
eSwatini	2016
Syria	1977
Tunisia	Before 2000
Trinidad and Tobago	2009
Türkiye	Before 2000
UK	Before 2000
Ukraine	2014
Vietnam	2020
Samoa	2018
South Africa	2000
Zimbabwe	2012

G Country comparisons

G.1 Cross-country results

Table G1: Covariate regressions for different export countries

	Dependent variable:			
	lnGap_USD			
	(SWE)	(FIN)	(NOR)	(DNK)
Duty Percentage	0.007***	0.008***	−0.001*	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
Democracy (1 = best)	−0.090***	−0.078***	−0.022	−0.165***
	(0.014)	(0.019)	(0.016)	(0.023)
Corruption (0 = best)	0.234***	0.168***	0.160***	0.341***
	(0.017)	(0.022)	(0.019)	(0.029)
Distance to capital (km)	−0.00001***	−0.00001***	−0.00002***	−0.00000***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
GDP	−0.000***	−0.000***	−0.000***	−0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
GDP per capita	0.00000***	−0.00000***	−0.00000***	0.00000***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Constant	0.116***	0.181***	0.491***	0.256***
	(0.014)	(0.018)	(0.016)	(0.023)
Observations	612,761	379,923	314,121	565,589
R^2	0.003	0.004	0.005	0.004
Adjusted R^2	0.003	0.004	0.005	0.004

Notes: Data are for Swedish, Finnish, Norwegian, and Danish exports to all countries, from 2000 - 2021, aggregated at the HS4 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of versions of Specification 2. The outcome variable is ln(Exports) - ln(Imports). The variables are as described in the notes to Table 3. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

Table G2: Fixed Effects specifications for different export countries

	Dependent variable:			
	lnGap_USD			
	(SWE)	(FIN)	(NOR)	(DNK)
Duty Percentage	0.009***	0.007***	0.005***	0.010***
	(0.001)	(0.001)	(0.001)	(0.001)
Democracy (1 = best)	−0.100*	−0.362***	−0.150***	−0.211**
	(0.052)	(0.068)	(0.058)	(0.092)
Corruption (0 = best)	0.232***	0.472***	−0.082	0.498***
	(0.060)	(0.074)	(0.069)	(0.106)
Country Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	612,761	379,923	314,121	565,589
R^2	0.038	0.030	0.063	0.044
Adjusted R^2	0.038	0.030	0.062	0.044

Notes: Data are for Swedish, Finnish, Norwegian, and Danish exports to all countries, from 2000 - 2021, aggregated at the HS4 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of Specification 3. The outcome variable is ln(Exports) - ln(Imports). The variables are as described in the notes to Table 3. The control variables not shown in the table are the GDP and GDP per capita. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

Table G3: Regressions with interaction terms for different export countries

	Dependent variable:			
	lnGap_USD			
	(SWE)	(FIN)	(NOR)	(DNK)
Duty Percentage	0.005***	0.009***	0.001	0.007***
	(0.001)	(0.001)	(0.001)	(0.002)
Duty x low democracy	0.005**	0.009***	0.008***	−0.004
	(0.002)	(0.003)	(0.002)	(0.004)
Duty x medium democracy	−0.001	0.0001	−0.001	0.004
	(0.002)	(0.002)	(0.002)	(0.003)
Low democracy	0.072***	0.162***	0.020	0.144***
	(0.024)	(0.030)	(0.027)	(0.039)
Medium democracy	0.084***	0.118***	0.084***	0.092***
	(0.016)	(0.020)	(0.017)	(0.026)
Duty x medium corruption	0.006***	−0.009***	0.004**	0.007**
	(0.002)	(0.002)	(0.002)	(0.003)
Duty x high corruption	0.0002	−0.004	−0.0004	0.004
	(0.003)	(0.003)	(0.003)	(0.004)
Medium corruption	−0.006	0.077***	−0.001	0.076**
	(0.018)	(0.023)	(0.019)	(0.036)
High corruption	0.005	0.050	0.021	0.203***
	(0.026)	(0.032)	(0.029)	(0.046)
Country Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
p-value: duty*lowDem = duty*medDem	<0.01	<0.01	<0.01	<0.01
p-value: duty*medCorr = duty*highCorr	<0.01	0.0131	0.0517	0.049
Observations	612,761	397,532	314,121	565,589
R^2	0.038	0.030	0.063	0.044
Adjusted R^2	0.038	0.029	0.062	0.044

Notes: Data are for Swedish, Finnish, Norwegian, and Danish exports to all countries, from 2000 - 2021, aggregated at the HS4 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of versions of Specification 4 for the different export countries. The outcome variable is ln(Exports) - ln(Imports). Tertiles for the democracy and corruption variables are recalculated on a yearly basis. The variables are as described in the notes to Table 3. The control variables not shown in the table are the GDP and GDP per capita. The p-values shown in the bottom of the table refer to a two-sided t-test verifying the equality of the coefficients on different interaction terms. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

G.2 Results for Finnish exports

Table G4: Finland - Regressions with different covariates

	Dependent variable:					
	lnGap_USD					
	(1)	(2)	(3)	(4)	(5)	(6)
Duty Percentage	0.012***	0.007***	0.005***	0.005***	0.008***	0.008***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Democracy (1 = best)		−0.265***		−0.020	−0.042**	−0.078***
		(0.013)		(0.019)	(0.019)	(0.019)
Corruption (0 = best)			0.362***	0.347***	0.337***	0.168***
			(0.013)	(0.019)	(0.019)	(0.022)
Distance to capital (km)					−0.00001***	−0.00001***
					(0.00000)	(0.00000)
GDP						−0.000***
						(0.000)
GDP per capita						−0.00000***
						(0.00000)
Constant	0.025***	0.211***	−0.033***	−0.016	0.055***	0.181***
	(0.003)	(0.010)	(0.004)	(0.016)	(0.016)	(0.018)
Observations	397,887	397,532	397,585	397,532	379,923	379,923
R^2	0.001	0.002	0.003	0.003	0.004	0.004
Adjusted R^2	0.001	0.002	0.003	0.003	0.004	0.004

Notes: Data are for Finnish exports to all countries, from 2000 - 2021, aggregated at the HS4 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of versions of Specification 2. The outcome variable is ln(Exports) - ln(Imports). The variables are as described in the notes to Table 3. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

Table G5: Finland - Regressions with different FTA use assumptions

	Dependent variable:				
	lnGap_USD				
	(1)	(2)	(3)	(4)	(5)
Duty - No FTAs	0.0005**				
	(0.0002)				
Duty - 25 percent FTA usage		0.001***			
		(0.0003)			
Duty - 50 percent FTA usage			0.002***		
			(0.001)		
Duty - 75 percent FTA usage				0.004***	
				(0.001)	
Duty - 100 percent FTA usage					0.008***
					(0.001)
Distance to capital (km)	−0.00001***	−0.00001***	−0.00001***	−0.00001***	−0.00001***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Democracy (1 = best)	−0.091***	−0.091***	−0.089***	−0.085***	−0.078***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Corruption (0 = best)	0.200***	0.197***	0.192***	0.180***	0.168***
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Constant	0.194***	0.193***	0.192***	0.188***	0.181***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	379,923	379,923	379,923	379,923	379,923
R^2	0.004	0.004	0.004	0.004	0.004
Adjusted R^2	0.004	0.004	0.004	0.004	0.004

Notes: Data are for Finnish exports to all countries, from 2000 - 2021, aggregated at the HS4 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of Specification 5. The outcome variable is ln(Exports) - ln(Imports). Column (5) presents our preferred version of specification, assuming 100% FTA usage. The variables are as described in the notes to Table 3. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

Table G6: Finland - Fixed Effects specifications

	Dependent variable:			
	lnGap_USD			
	(6)	(7)	(8)	(9)
Duty Percentage	0.008***	0.008***	0.007***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
Democracy (1 = best)	−0.078***	−0.105	−0.101***	−0.362***
	(0.019)	(0.064)	(0.019)	(0.068)
Corruption (0 = best)	0.168***	0.667***	0.180***	0.472***
	(0.022)	(0.072)	(0.022)	(0.074)
Distance to capital (km)	−0.00001***		−0.00001***	
	(0.00000)		(0.00000)	
Constant	0.181***			
	(0.018)			
Country Fixed Effects	No	Yes	No	Yes
Year Fixed Effects	No	No	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	379,923	379,923	379,923	379,923
R^2	0.004	0.030	0.006	0.030
Adjusted R^2	0.004	0.029	0.006	0.030

Notes: Data are for Finnish exports to all countries, from 2000 - 2021, aggregated at the HS4 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of Specifications 2 and 3. The outcome variable is ln(Exports) - ln(Imports). Column (6) presents our preferred version of Specification 2, i.e. with all the covariates included, but with no Fixed Effects. Columns (7) to (9) add Country FE, Year FE, and then both. The variables are as described in the notes to Table 3. The control variables not shown in the table are the GDP and GDP per capita. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

Table G7: Finland - Regressions with interaction terms

	Dependent variable:		
	lnGap_USD		
	(10)	(11)	(12)
Duty Percentage	0.007***	0.008***	0.008***
	(0.001)	(0.001)	(0.001)
Duty x low democracy	0.004**		0.010***
	(0.002)		(0.003)
Duty x medium democracy	−0.004**		0.003
	(0.002)		(0.003)
Low democracy	0.089***		0.070**
	(0.032)		(0.033)
Medium democracy	0.101***		0.087***
	(0.019)		(0.019)
Duty x medium corruption		−0.005***	−0.011***
		(0.002)	(0.003)
Duty x high corruption		0.003	−0.004
		(0.002)	(0.003)
Medium corruption		0.104***	0.095***
		(0.023)	(0.023)
High corruption		0.084***	0.073**
		(0.031)	(0.032)
Country Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Controls	Yes	Yes	Yes
p-value: duty*lowDem = duty*medDem	<0.01		<0.01
p-value: duty*medCorr = duty*highCorr		<0.01	<0.01
Observations	397,585	397,585	397,585
R^2	0.029	0.029	0.030
Adjusted R^2	0.029	0.029	0.029

Notes: Data are for Finnish exports to all countries, from 2000 - 2021, aggregated at the HS4 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of versions of Specification 4 with different interaction terms included. The outcome variable is ln(Exports) - ln(Imports). Column (12), our preferred specification, includes interaction terms between the duty percentage and tertiles of both the democracy and corruption scores. Tertiles for the democracy and corruption variables are recalculated on a yearly basis. The variables are as described in the notes to Table 3. The control variables not shown in the table are the GDP and GDP per capita. The p-values shown in the bottom of the table refer to a two-sided t-test verifying the equality of the coefficients on different interaction terms. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

G.3 Results for Danish exports

Table G8: Denmark - Regressions with different covariates

	Dependent variable:					
	lnGap_USD					
	(1)	(2)	(3)	(4)	(5)	(6)
Duty Percentage	0.009***	0.006***	0.006***	0.005***	0.006***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Democracy (1 = best)		−0.278***		−0.138***	−0.171***	−0.165***
		(0.011)		(0.016)	(0.016)	(0.016)
Corruption (0 = best)			0.303***	0.200***	0.200***	0.341***
			(0.012)	(0.017)	(0.017)	(0.019)
Distance to capital (km)					−0.00001***	−0.00000***
					(0.00000)	(0.00000)
GDP						−0.000***
						(0.000)
GDP per capita						0.00000***
						(0.00000)
Constant	0.258***	0.449***	0.206***	0.318***	0.364***	0.256***
	(0.003)	(0.008)	(0.003)	(0.013)	(0.014)	(0.016)
Observations	566,667	565,589	565,769	565,589	565,589	565,589
R^2	0.001	0.002	0.002	0.002	0.003	0.004
Adjusted R^2	0.001	0.002	0.002	0.002	0.003	0.004

Notes: Data are for Danish exports to all countries, from 2000 - 2021, aggregated at the HS4 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of versions of Specification 2. The outcome variable is ln(Exports) - ln(Imports). The variables are as described in the notes to Table 3. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

Table G9: Denmark - Regressions with different FTA use assumptions

	Dependent variable:				
	lnGap_USD				
	(1)	(2)	(3)	(4)	(5)
Duty - No FTAs	0.002***				
	(0.0004)				
Duty - 25 percent FTA usage		0.003***			
		(0.001)			
Duty - 50 percent FTA usage			0.004***		
			(0.001)		
Duty - 75 percent FTA usage				0.007***	
				(0.001)	
Duty - 100 percent FTA usage					0.007***
					(0.001)
Distance to capital (km)	−0.00000***	−0.00000***	−0.00000***	−0.00000***	−0.00000***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Democracy (1 = best)	−0.175***	−0.173***	−0.171***	−0.167***	−0.165***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Corruption (0 = best)	0.362***	0.357***	0.349***	0.338***	0.341***
	(0.019)	(0.020)	(0.020)	(0.019)	(0.019)
Constant	0.265***	0.264***	0.262***	0.258***	0.256***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	565,589	565,589	565,589	565,589	565,589
R^2	0.004	0.004	0.004	0.004	0.004
Adjusted R^2	0.004	0.004	0.004	0.004	0.004

Notes: Data are for Danish exports to all countries, from 2000 - 2021, aggregated at the HS4 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of Specification 5. The outcome variable is ln(Exports) - ln(Imports). Column (5) presents our preferred version of specification, assuming 100% FTA usage. The variables are as described in the notes to Table 3. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

Table G10: Denmark - Fixed Effects specifications

	Dependent variable:			
	lnGap_USD			
	(6)	(7)	(8)	(9)
Duty Percentage	0.007***	0.009***	0.009***	0.010***
	(0.001)	(0.001)	(0.001)	(0.001)
Democracy (1 = best)	−0.165***	−0.589***	−0.126***	−0.211***
	(0.016)	(0.054)	(0.016)	(0.058)
Corruption (0 = best)	0.341***	0.228***	0.266***	0.498***
	(0.019)	(0.068)	(0.020)	(0.069)
Distance to capital (km)	−0.00000***		−0.00000***	
	(0.00000)		(0.00000)	
Constant	0.256***			
	(0.016)			
Country Fixed Effects	No	Yes	No	Yes
Year Fixed Effects	No	No	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	565,589	565,589	565,589	565,589
R^2	0.004	0.043	0.028	0.044
Adjusted R^2	0.004	0.042	0.028	0.044

Notes: Data are for Danish exports to all countries, from 2000 - 2021, aggregated at the HS4 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of Specifications 2 and 3. The outcome variable is ln(Exports) - ln(Imports). Column (6) presents our preferred version of Specification 2, i.e. with all the covariates included, but with no Fixed Effects. Columns (7) to (9) add Country FE, Year FE, and then both. The variables are as described in the notes to Table 3. The control variables not shown in the table are the GDP and GDP per capita. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

Table G11: Denmark - Regressions with interaction terms

	Dependent variable:		
	lnGap_USD		
	(10)	(11)	(12)
Duty Percentage	0.008***	0.007***	0.007***
	(0.001)	(0.001)	(0.001)
Duty x low democracy	0.001		−0.002
	(0.002)		(0.002)
Duty x medium democracy	0.008***		0.005**
	(0.001)		(0.002)
Low democracy	0.172***		0.138***
	(0.030)		(0.029)
Medium democracy	0.075***		0.064***
	(0.016)		(0.016)
Duty x medium corruption		0.007***	0.006***
		(0.001)	(0.002)
Duty x high corruption		0.002	0.003
		(0.003)	(0.003)
Medium corruption		0.091***	0.083***
		(0.019)	(0.019)
High corruption		0.232***	0.208***
		(0.029)	(0.029)
Country Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Controls	Yes	Yes	Yes
p-value: duty*lowDem = duty*medDem	<0.01		<0.01
p-value: duty*medCorr = duty*highCorr		<0.01	0.0135
Observations	565,769	565,769	565,769
R^2	0.044	0.044	0.044
Adjusted R^2	0.044	0.044	0.044

Notes: Data are for Danish exports to all countries, from 2000 - 2021, aggregated at the HS4 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of versions of Specification 4 with different interaction terms included. The outcome variable is ln(Exports) - ln(Imports). Column (12), our preferred specification, includes interaction terms between the duty percentage and tertiles of both the democracy and corruption scores. Tertiles for the democracy and corruption variables are recalculated on a yearly basis. The variables are as described in the notes to Table 3. The control variables not shown in the table are the GDP and GDP per capita. The p-values shown in the bottom of the table refer to a two-sided t-test verifying the equality of the coefficients on different interaction terms. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

G.4 Results for Norwegian exports

Table G12: Norway - Regressions with different covariates

	Dependent variable:					
	lnGap_USD					
	(1)	(2)	(3)	(4)	(5)	(6)
Duty Percentage	−0.001*	−0.003***	−0.008***	−0.007***	−0.002***	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Democracy (1 = best)		−0.098***		0.144***	0.012	−0.022
		(0.016)		(0.023)	(0.023)	(0.023)
Corruption (0 = best)			0.276***	0.385***	0.370***	0.160***
			(0.018)	(0.025)	(0.025)	(0.029)
Distance to capital (km)					−0.00002***	−0.00002***
					(0.00000)	(0.00000)
GDP						−0.000***
						(0.000)
GDP per capita						−0.00000***
						(0.00000)
Constant	0.322***	0.393***	0.288***	0.171***	0.343***	0.491***
	(0.004)	(0.012)	(0.004)	(0.019)	(0.021)	(0.023)
Observations	314,647	314,121	314,183	314,121	314,121	314,121
R^2	0.00001	0.0001	0.001	0.001	0.003	0.005
Adjusted R^2	0.00001	0.0001	0.001	0.001	0.003	0.005

Notes: Data are for Norwegian exports to all countries, from 2000 - 2021, aggregated at the HS4 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of versions of Specification 2. The outcome variable is ln(Exports) - ln(Imports). The variables are as described in the notes to Table 3. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

Table G13: Norway - Regressions with different FTA use assumptions

	Dependent variable:				
	lnGap_USD				
	(1)	(2)	(3)	(4)	(5)
Duty - No FTAs	0.003***				
	(0.001)				
Duty - 25 percent FTA usage		0.003***			
		(0.001)			
Duty - 50 percent FTA usage			0.002*		
			(0.001)		
Duty - 75 percent FTA usage				0.0002	
				(0.001)	
Duty - 100 percent FTA usage					−0.001
					(0.001)
Distance to capital (km)	−0.00002***	−0.00002***	−0.00002***	−0.00002***	−0.00002***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Democracy (1 = best)	−0.021	−0.021	−0.021	−0.022	−0.022
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Corruption (0 = best)	0.125***	0.131***	0.139***	0.150***	0.160***
	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)
Constant	0.489***	0.489***	0.490***	0.491***	0.491***
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	314,121	314,121	314,121	314,121	314,121
R^2	0.005	0.005	0.005	0.005	0.005
Adjusted R^2	0.005	0.005	0.005	0.005	0.005

Notes: Data are for Norwegian exports to all countries, from 2000 - 2021, aggregated at the HS4 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of Specification 5. The outcome variable is ln(Exports) - ln(Imports). Column (5) presents our preferred version of specification, assuming 100% FTA usage. The variables are as described in the notes to Table 3. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

Table G14: Norway - Fixed Effects specifications

	Dependent variable:			
	lnGap_USD			
	(6)	(7)	(8)	(9)
Duty Percentage	−0.001	0.003***	0.002**	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)
Democracy (1 = best)	−0.022	−0.384***	0.009	−0.150
	(0.023)	(0.085)	(0.024)	(0.092)
Corruption (0 = best)	0.160***	−0.280***	−0.079**	−0.082
	(0.029)	(0.104)	(0.031)	(0.106)
Distance to capital (km)	−0.00002***		−0.00002***	
	(0.00000)		(0.00000)	
Constant	0.491***			
	(0.023)			
Country Fixed Effects	No	Yes	No	Yes
Year Fixed Effects	No	No	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	314,121	314,121	314,121	314,121
R^2	0.005	0.062	0.035	0.063
Adjusted R^2	0.005	0.061	0.035	0.062

Notes: Data are for Norwegian exports to all countries, from 2000 - 2021, aggregated at the HS4 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of Specifications 2 and 3. The outcome variable is ln(Exports) - ln(Imports). Column (6) presents our preferred version of Specification 2, i.e. with all the covariates included, but with no Fixed Effects. Columns (7) to (9) add Country FE, Year FE, and then both. The variables are as described in the notes to Table 3. The control variables not shown in the table are the GDP and GDP per capita. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

Table G15: Norway - Regressions with interaction terms

	Dependent variable:		
	lnGap_USD		
	(10)	(11)	(12)
Duty Percentage	0.003*	0.001	0.002
	(0.002)	(0.002)	(0.002)
Duty x low democracy	0.008***		0.007*
	(0.002)		(0.004)
Duty x medium democracy	−0.001		−0.003
	(0.002)		(0.003)
Low democracy	0.0001		0.001
	(0.043)		(0.045)
Medium democracy	0.096***		0.098***
	(0.024)		(0.025)
Duty x medium corruption		0.006***	0.005
		(0.002)	(0.003)
Duty x high corruption		0.005*	0.0004
		(0.003)	(0.004)
Medium corruption		0.003	0.003
		(0.035)	(0.036)
High corruption		0.020	0.019
		(0.045)	(0.047)
Country Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Controls	Yes	Yes	Yes
p-value: duty*lowDem = duty*medDem	<0.01		<0.01
p-value: duty*medCorr = duty*highCorr		0.4203	0.079
Observations	314,183	314,183	314,183
R^2	0.063	0.063	0.063
Adjusted R^2	0.062	0.062	0.062

Notes: Data are for Norwegian exports to all countries, from 2000 - 2021, aggregated at the HS4 level of the harmonised system as explained in Section 5.3.1. Robust standard errors reported in parentheses. This table reports the results of versions of Specification 4 with different interaction terms included. The outcome variable is ln(Exports) - ln(Imports). Column (12), our preferred specification, includes interaction terms between the duty percentage and tertiles of both the democracy and corruption scores. Tertiles for the democracy and corruption variables are recalculated on a yearly basis. The variables are as described in the notes to Table 3. The control variables not shown in the table are the GDP and GDP per capita. The p-values shown in the bottom of the table refer to a two-sided t-test verifying the equality of the coefficients on different interaction terms. The levels of significance are *p<0.1, **p<0.05, and ***p<0.01.

H Event study

In order to further study our hypothesis that political institutions play an important role in customs fraud, we conduct an event study focusing on coups and other irregular changes in leadership. We use the Archigos dataset for this analysis, which contains information on coups and other irregular government transitions, as well as country leader characteristics (Goemans et al., 2009).

The empirical specification is as follows:

$$y_{ijk} = \sum_{t \neq 0}^{T} \beta_t . \boldsymbol{I}[t = k] + X + \zeta_i + \gamma_j + \epsilon_{ijk} \tag{6}$$

Where: $\boldsymbol{I}$ is a period indicator, X is a vector of the covariates found in Equation 2, ζ_i is import country fixed effects, and γ_j is year fixed effects.

To test our intuition that coups or other irregular government transitions will increase log trade gaps, we propose to conduct an event study utilising a subsample of countries that had an irregular change in power. In order to have at least five years of data before and seven years after the event, we restrict our sample to irregular changes which occurred between 2002 and 2018. We consider the year during which the change of power happened as period 0 and the rest of the periods are indexed relative to it. The coefficients of interest are the sequence β_t which will give us the size of the effect relative to period 0. We suspect that the coefficients on β_t will not be significantly different from zero in periods before t = 0, and that they will be positive when t > 0.

Based on our previous findings, we would expect to see a significant increase in trade gaps after irregular government transitions as these are usually situations in which democracy scores decrease and opportunities for increased corruption abound. To analyse these effects, we reduce our sample to countries which experienced only one such transition in the years between 2005 and 2014. We discard countries which had multiple transitions as it is not possible to disentangle the effects of one episode from another. We follow the definition of "irregular" used by the Archigos dataset, in that a transition is irregular if it does not follow established rules and conventions. This could be due to a revolt or coup, or an assassination of a sitting leader (Goemans et al., 2009).

Our reduced sample has twelve countries[4]. Due to the small number of countries, we present our results at the HS2 level, instead of the HS4 level. This is mainly because the sample is skewed due to significant exports to countries like Egypt and Thailand, but many fewer exports to countries like Fiji. At lower levels of granularity, the sample is significantly more balanced.

[4]The countries are: Burkina Faso, Central African Republic, Côte d'Ivoire, Egypt, Fiji, Guinea, Honduras, Libya, Madagascar, Mali, Niger, and Thailand.

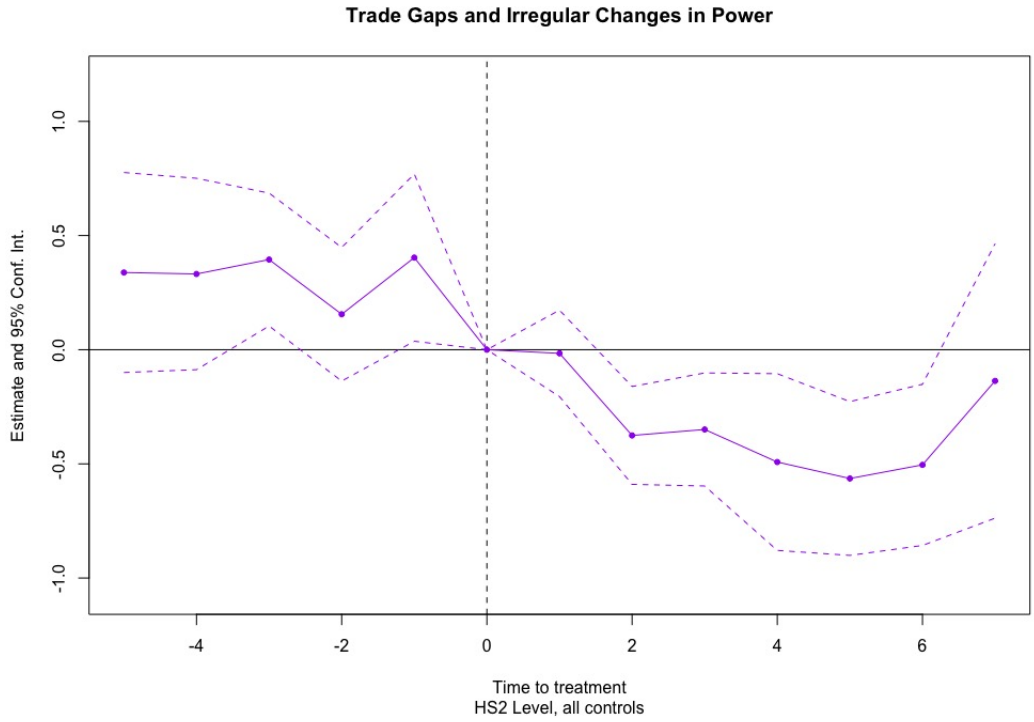

Figure H1: Trade gaps relative to irregular changes in power, aggregated at the HS2 level

Contrary to our expectations, we find (as shown in Figure H1) that irregular government transitions actually *reduce* the incidence of undervaluation fraud. While trade gaps before the transition are relatively stable in the sample, there is a clear decrease in the years after it. Specifically, in the years between two and six years after an irregular transition, we find that trade gaps are around 50% lower than in the year the irregular transition occurred. We propose two possible mechanisms for this. First, it is reasonable to expect that just after a coup, the incoming government is keen to raise funds and extract as many rents as possible from the state. Therefore increased scrutiny on customs in order to increase tax revenues is likely. Secondly, during irregular transitions functionaries and officers in the customs and tax administrations could get replaced, and then relationships through which corruption is carried out get broken. It takes some time for these to be rebuilt, so this could explain the gradual return to zero. We accept however that due to the lack of specific literature on this topic, these explanations remain conjecture for the time being.