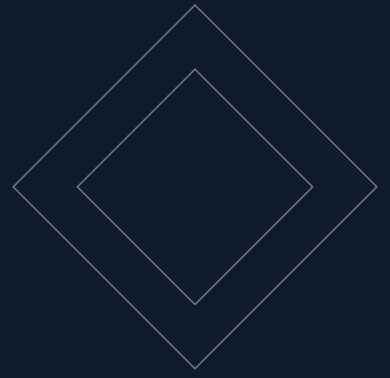




---

REBECCANOMICS · RESEARCH PAPER

# Quantitative Analysis of the Move to Paperless Trade

---

July 2026



A STUDY BY

**Dr Rebecca Harding**

Independent Economist · Rebeccanomics

**Rebeccanomics**

# Quantitative Analysis of the Move to Paperless Trade

Dr Rebecca Harding — July 2026

---

## Abstract

---

There are clear economic and political imperatives to accelerate the process of legal reform to enable paperless trade that emerge from the research conducted for this report. At present, the trend forecasts for trade across the Commonwealth to 2026 suggest that on average trade will increase by around 0.2 per cent annually. This is a potential trajectory that will not help to allow Commonwealth nations, especially in emerging countries, to recover from the pandemic quickly enough through trade.

The analysis for this report was based on a summary methodology as follows:

Public discourse analysis covering 600 articles and over 1.1 m words across the Commonwealth.

Documentary and literature search.

Twenty qualitative semi-structured interviews with key stakeholders and practitioners in trade.

Quantitative forecasting of impact of regulatory reform by using Bayesian forecasting methodologies and based on Ease of Doing Business survey data, alongside a survey of 55 international banks and the Coriolis Technologies MultiLateral trade data covering Commonwealth trade.

The core findings are that digital trade facilitation across borders could increase trade across the Commonwealth by around US\$90 billion, and legal reform to support the digitalisation of electronic records could unleash as much as US\$1.1 trillion, bringing the total benefits from paperless trade to nearly US\$1.2 trillion by 2026. More than this, the analysis suggests a reduction in Commonwealth exporters' cost base of around 75 per cent on average, which compensates for the excessive costs relative to revenues currently experienced by some Island and African economies.

To be clear, these figures are aspirational and based on the assumption that costs will fall and trade increase from a shift to paperless trade enabled by legal reform in the next 18 months. They should be seen as a target rather than as a definitive statement of what will happen but are nevertheless a strong business case for implementing changes at the earliest opportunity.

## Introduction

---

This report aims to quantify the potential impact of legal reform to enable the use of the so-called "transferable records" on Commonwealth trade. "Transferable records" are paper-based documents or instruments used in domestic or international trade and trade finance such as bills of lading, bills of exchange, promissory notes, warehouse receipts, guarantees and standby letters of credit. Since the beginnings of trade between individuals, companies and nations, these records have been manual; there are an estimated 4 billion paper-based documents that are being processed at any one point in time around the world according to the International Chamber of Commerce (ICC).

The Model Law on Electronic Transferrable Records (MLETR) aims to enable "the legal use of transferable records both domestically and across borders". It is the United Nations Commission on International Trade (UNCITRAL)'s legal framework for enabling digital transferable records in law across the UN membership. The goal is to facilitate trade speeding up processing and transmission times, making them more secure, automated and usable in multiple settings.

Clearly, anything that makes the process of processing documents more quickly is to be welcomed. Estimates vary on how long it takes to process manual documents, but, according to the Boston Consulting Group

(BCG), 60 per cent of businesses and 61 per cent of banks see paper documentation as a major pain point in trade finance. Similarly, the ICC UK survey of 55 international banks suggested that processing documents took on average per trade finance deal around 19 days, or just under 4 weeks, but that trade finance solutions, such as receivables finance, letters of credit and working capital solutions, take more than 4 weeks to process.

More than this, a quick calculation serves to point out the importance of this market in total trade terms: if globally there are 4 billion transactions on average per year, if global trade was worth around US\$21.6 billion in 2019 and if there are an average of 15 documents per transaction, then on average in 2019 the cost per transaction could be estimated at around US\$81,000. The consequence is that any transaction worth less than that value makes a loss for bank. As many exporters are smaller, especially across emerging markets and in Island economies, this very simple estimation shows why they are excluded from trade finance.

So the imperative for change is apparent generally and not just across the Commonwealth. Quite without the impact of the Covid pandemic and the need to accelerate the recovery from it, there would have been a business case. Technology is moving quickly and, as will be illustrated in this report, there is evidence that the appetite for digital and technology-based solutions in the public discourse is there. The challenge now is to harness this momentum with a political will to implement legal reform.

## **Aims and Objectives**

This report is a first step towards that goal. Its overall purpose is to undertake a quantitative analysis of adopting paperless trade documentation to national and regional economies across the Commonwealth, focusing on Less Developed Countries (LDCs) and Small Island Developing States (SIDS) in Africa, Asia, the Caribbean and the Pacific Islands. For this reason, the "larger" and/or more developed trading economies such as the UK, Australia, New Zealand, Singapore, Canada and South Africa have been treated separately in charts and not analysed in detail.

It is worth noting that Singapore has already adopted MLETR and that its activities in the space are seen as providing an excellent test case, indeed template, for action by other Commonwealth nations. The UK's appears to be moving towards adoption through amendments to its 1855 Bill of Lading Act that would make electronic records equal in law to paper ones. For the purposes of this paper, however, these economies are not regarded as starting at a different point in time in terms of the impact of legislative reform since actual digitally enabled trade under the new regime has only just started to be implemented in Singapore and will still take some time in the UK.

The specific objectives of the research therefore have been to:

- provide an overview of the current challenges facing Commonwealth nations in implementing paperless trade by means of qualitative desk research, discourse analysis and semi-structured in-depth interviews;
- assess the current state of trade across the Commonwealth and prospects for growth post-pandemic with a specific focus on less developed economies and island economies in Africa, Asia, the Caribbean and the Pacific;
- understand the specific costs that exporters face by means of interpolation from published data on the costs of exporting from the World Bank's Ease of Doing Business Survey focused specifically on the costs of exporting;
- assess how those costs can be reduced through implementation of paperless trade and broader digitisation processes – for example, at borders through Single Windows and in payments and trade finance through MLETR alignment – using the estimations from documentary research meta-analysis, discourse analysis and estimates from interviews of potential cost reductions as a result of streamlined processes;
- assess the impact on trade of (a) the reduction in costs for businesses, (b) the reduction in costs for banks and (c) the increase in revenues for businesses and (d) the increase in revenues for banks from new market creation. These estimates are derived similarly from the ICC survey of 55 international banks based in London and the

documentary, discourse and interview research conducted above. They provide the coefficients for trade growth, which are applied to the trade forecasting methodology described in the Appendix.

The research only aims to provide a business case in support of the Commonwealth's Connectivity Cluster Agenda that targets digital trade facilitation supported through an appropriate legislative framework and aims to create an increase in Commonwealth trade by US\$2 trillion in 2030. This agenda is critical to Commonwealth nations, particularly in emerging countries and island economies not just to streamline trade but also to enable trade-based solutions to post-pandemic recovery.

## **Research caveats**

Three caveats to the research conducted here need to be highlighted at the outset. First, there is no analysis of intra-Commonwealth trade. While this is doable, and intrinsically of interest, the Commonwealth is not a Free Trade Area and therefore the issue of facilitating trade within the group is to a large extent a point of debate. Instead, each individual country within the Commonwealth has been studied in isolation. To avoid double-counting, the analysis only looks at exports from each.

Second, there is no set of recommendations made from the research. The terms of reference for the research were simply to provide a quantitative assessment of the potential impact of the implementation uniformly of legal reform to enable electronic records across the Commonwealth. This, in itself, is a theoretical limitation to the research – it is unlikely that all countries would simultaneously introduce reforms as this would involve a degree of political alignment across the Commonwealth that would be challenging given that it is not a trading bloc.

Finally, and for the reasons above, it is important to re-emphasise that the quantifications provided in this research are based on assumptions about how costs would be reduced and trade increase as a result of streamlining and market creation, particularly enabling micro, small and medium-sized enterprises (MSMEs) to enter the market as a result of the changes enabled through greater digitisation. These are derived from qualitative research and are intrinsically "untestable" for their accuracy since no precedents exist. This is discussed in more detail in the methodology section and the detailed forecasting methodology is described in the Statistical Appendix.

Even with these limitations, the research presented here suggests that digital trade facilitation across borders could increase trade across the Commonwealth by around US\$90 billion, and legal reform to support the digitalisation of electronic records could unleash as much as US\$818 billion, bringing the total benefits from paperless trade to over US\$1 trillion by 2026 and thereby making a substantial contribution to the aspiration of a US\$2 trillion increase in trade across the Commonwealth by 2030. More than this, the analysis suggests a reduction in Commonwealth exporters' cost base of around 75 per cent on average, which compensates for the excessive costs relative to revenues currently experienced by some Island and African economies.

## **Report structure**

This report is structured as follows:

Section 1 looks at MLETR and provides some "plain English" terms as they are used in this paper. The concept is a simple one – that documentation used in trade that has hitherto been paper-based can be digitised. In so doing, there is scope for speeding up their scrutiny and approval thereby reducing the potential for compliance fraud, reducing the costs of delay for businesses generally and increasing access for MSMEs in particular.

Section 2 is a summary of the research methodology undertaken and should be read in conjunction with the Statistical Appendix. It highlights the fact that assessing the impact of legal reform to enable electronic transferable records is intrinsically difficult since there are no empirical precedents. This research and its method should be seen as a first attempt and a starting point along that journey as a result.

Section 3 provides some background on international trade across the Commonwealth and examines the issues that each of the regions covered in this research faces, particularly as they start to recover from the Covid pandemic. One important factor is drawn out of this discussion that underpins the approach to the quantitative analysis undertaken. Many of the economies in the Commonwealth have specific challenges because they are under-developed or because they are island economies. This means that while legal reform is important as a catalyst for trade facilitation, there are other issues such as reform to border crossings, inclusion and enabling infrastructures that need to be addressed alongside an enabling digital framework.

Section 4 presents the findings of the research. These are presented for each country in the Commonwealth and grouped into specific regions as follows: the "larger developed" economies, Africa, Asia and the Caribbean, highlighting the status of island economies in the Pacific and the Caribbean. The most striking finding is that at present, the average costs of international trade for some island economies and some nations in Africa are higher than the average revenues received. This makes trade untenable in those countries and highlights the importance of research like this to enable progress towards workable solutions.

Section 5 provides some concluding remarks on the research itself and the call for action it represents. The case is clearly made for legal reform to enable electronic transferrable records and for the implementation of digital trade, such as Single Windows, Digital IDs and digital payments, not just to boost trade generally but also to create a more inclusive and sustainable recovery from the Covid pandemic. However, it also highlights the need for education and trust building as a means of making sure that there is acceptance of new digital trade systems to avoid the risk that lack of trust undermines their take-up once implemented.

As a note of context, one thing is important and came out from all the interviews, documentary research and discussions held during the course of this research: governments have a strategic role to play as regulators and can use this role to facilitate trade in a relatively low-cost way. It does not take away the need for long-term investment in infrastructure or other trade facilitation enablers in the business environment. But regulatory reform does signal an intent to address structural issues at a national and international level. Here the Commonwealth has a role to play alongside international organisations such as the International Chamber of Commerce and the United Nations Development Programme in providing enabling frameworks that make digitalisation the trigger for enabling equal access for MSMEs and large businesses alike to the benefits of trade-based growth.

## Section 1: Background: MLETR and terminology as it is used in this research

---

It is worth spending some time on understanding precisely the terminology as it will be used in this report. MLETR, as described above, is the Model Law to enable Electronic Transferable Records in trade: that is, a legal framework provided by the UNCITRAL which has been adopted by the International Chamber of Commerce (ICC) Digital Standards Initiative (DSI) to provide the support to governments and businesses to speed up the process of digital trade adoption.

In lay terms, and at the risk of sounding repetitive, this is the process of moving from paper-based trade documentation to electronic trade documentation. Technology exists to replace paper, through the use of digital technologies such as blockchain or cloud-based solutions, including machine learning and artificial intelligence. It is seen as having four primary advantages:

**Increasing access to finance:** either through bank-based or through Fintech-based solutions it enables trade finance "assets" (financial documents such as Bills of Lading, Bills of Exchange or Promissory Notes) to be made available or "distributed" across banks, institutional investors or non-bank trade finance providers (such as invoice finance or supply chain finance providers). This increases the amount of available finance for SMEs. This advantage comes simply from the fact that the documents themselves are instruments through which finance can be enabled. Making them digital means that they go to a wider potential market in a secure way.

**Reducing costs:** Digital documentation simply speeds up the processing time. If everything is manual, then estimates by the ICC, and highlighted above, suggest that it can take more than 4 weeks to process the documentation before awarding finance to a particular transaction or project. This is a burden for the SME and for the trade finance provider.

**Reducing the risk of fraud:** the main sources of fraud in current systems are multiple uses of one document for financing purposes, the lack of transparency in the trade system, which means that documents can potentially be used for financial or insurance purposes that are not genuine or not the original, and the misallocation of product codes to avoid taxation or identification as a dual-use good. This risk is likely to become greater as we move towards requirements for greater transparency on environment, social and governance (ESG) grounds.

**Creating global standards:** there are currently no common standards in global trade outside of Customs and Excise frameworks, dual-use goods and prohibited trade, and increasingly sustainable financial disclosures (SFD). The process of digitisation has not only the capacity to accelerate the process of standardisation in global trade but is also a pre-requisite of the digitisation process itself in order to avoid further complexity.

The ICC is working with organisations such as the Commonwealth to achieve standardisation, interoperability (in other words, shared rules internationally), legislative reform and common technical standards as well as work to achieve international trade standardisation starting with the documentation required around Bills of Lading for containers and dry bulk where currently the reporting requirements are opaque.

Bills of Lading are the logical starting point. A Bill of Lading is the legal document issued by the carrier that attributes the type and quantity of goods shipped, their port of origin and their destination. It acts like a receipt for the shipper confirming that something has been put on board a ship and is destined for another country and a buyer. As such, it confirms that something has been produced and shipped and can therefore be used as a means to leverage working capital by the shipper from when it leaves the port to when it is picked up and approved by the buyer at the other end. The majority of these processes are currently manual and non-standard causing delays, leaving the system open to fraud and creating inconsistencies before at and the port itself.

There are two ways in which this is important for the Commonwealth's LDCs and SIDS. First, the non-standardised and manual nature of a Bill of Lading makes border processes complex and costly for the exporter. Lower levels of education or literacy mean that reams of manual documentation in customs and excise forms, identification requirements, rules of origin and Incoterms compliance present a major challenge for some groups, and especially women in Africa. Produce may deteriorate during this weighty compliance process at the border, but more than that it leaves the processes wide open to abuse by officials who may require "informal" payments to speed the border process up. This was found by our research to be a major challenge for all exporters in Africa and is corroborated by the very high costs of exporting seen in the Ease of Doing Business surveys and illustrated in Figure 1.

*Figure 1. Average pre-border and border costs for Commonwealth countries, 2019. Source: Author's calculations based on documentary research, World Bank Ease of Doing Business survey 2020 and Transparency International Corruption Indicators. The chart breaks down costs per US\$25,000 unit into border compliance, documentary, transportation and informal costs.*

The chart serves to point out that the informal costs are the most significant aspect of pre-border and border crossing costs in all LDC and SIDS economies. In contrast, they are zero in the developed economies.

Second, the Bill of Lading document, if standardised and digitised, could potentially represent a way around the challenges faced by LDCs and SIDS generally and in the Commonwealth in particular by creating a document that is authenticated and immutable in any jurisdiction, irrespective of any broader compliance (Anti Money Laundering or Know Your Client) risks associated with trade in emerging markets. This is important because the trade finance gap has been widening during the Covid Pandemic and many smaller businesses in emerging economies are unable to access finance.

The trade finance gap is estimated by the Asian Development Bank to have increased to US\$1.7 trillion during the pandemic with around 40 per cent of applications by SMEs for trade finance rejected. The common causes of the trade finance gap cited by many are seen as being the lack of correspondent banks, the high risks of fraud, the costs of due diligence relative to the size of any specific trade finance deal, and the lack of trade credit insurance and buyer information for emerging markets, especially Africa. During the pandemic, emerging market banks across the globe have been concerned about demand for their customers' exports (71 per cent), refinancing needs (71 per cent) and challenges with operational cashflow (70 per cent) over all but with East Asia and the Pacific and South Asia likely to be most hit by the drop in demand for exports (86 and 94 per cent, respectively) and sub-Saharan Africa most at risk from supply chain disruptions. The risks to SIDS is not just from the drop in physical trade, but also in the collapse of tourism and other service-related trade.

Clearly, not all of these Covid-related issues are addressable through digitisation of Bills of Lading, but even for services, digital trade documentation (buyer and seller digital Identification, and digital service contracts and invoices, for example) will speed up and make more secure the process of accessing trade finance both between a buyer and a seller, but also through supply chains which have suffered significantly as a result of the pandemic. Digitisation is often also seen as a route to addressing the issue of risk related to fraud, AML and KYC which, as the costs associated with the informal economy highlighted in Figure 1 suggest, are endemic in emerging markets.

In summary, there are two areas where electronic transferable records or digital documentation can have an impact:

- as a means of reducing the costs of trade and
- as a mechanism for enabling finance.

The two are clearly related to one another, but the important point here is that if the costs can be reduced, especially informal costs, this has a multiplier effect on the exports of a country because it helps with a broader trade facilitation agenda and, critically, enables streamlined access to finance for smaller and riskier deals which are the characteristic of emerging markets. The research reported in this paper is an attempt to quantify each.

## Section 2: A note on methodology

---

The approach taken in this research seeks to address a number of methodological challenges inherent to quantifying what is intrinsically unknowable prior to implementation. These are:

- no direct data on the size of the digital trade finance market anywhere, still less the Commonwealth LDCs and SIDS;
- bill of Lading data limited to goods data and therefore not giving a full picture of Commonwealth international businesses;
- little or no quantitative research into the role of digitisation in enabling SME trade finance growth or trade growth therefore limited scope for deriving an appropriate methodology from the literature.

This means that a variety of data sources have been used in combination to estimate both the current state of trade and progress towards digitalisation and the prospects of cost reductions and increased access to finance in the next 5 years. A 5-year period of time has been taken as the benchmark because of the urgency of catalysing recovery from the current Covid pandemic. These sources are:

- Twenty semi-structured interviews with trade practitioners, banks and experts based in the Commonwealth's LDCs and SIDS or with a view of trade in those nations;

Coriolis Technologies trade data covering intra and extra Commonwealth trade in goods and services as well as individual country-level trade for each Commonwealth nation;

Coriolis Technologies Bill of Lading data covering marine shipments to provide recent impact assessment of Commonwealth countries where data are available;

World Bank Ease of Doing business survey data;

Transparency International Corruption Indices;

UNCTAD freight cost data;

Survey of 55 international banks and 20 international businesses based in London conducted for the ICC UK in Q1 2021 but with estimates of time and cost savings and revenue growth;

Discourse analysis: python-based web-scraping techniques across 600 media outlets in total across the Commonwealth and the African continent.

## Research questions and approach

The research sought to answer four questions:

**Are there any barriers to the implementation of paperless trade across the Commonwealth nations?** This question is addressed by means of a discourse analysis covering 100 media outlets across the Commonwealth, semi-structured interviews and desk research.

**What are the current biggest challenges facing Commonwealth nations in trade terms?** This question is addressed by means of discourse analysis of additional 500 media sources largely in Africa and Coriolis Technologies trade and Bill of Lading data.

**What are the costs faced by Commonwealth LDCs and SIDS currently?** The costs for the current year are estimated from the Ease of Doing Business survey for pre-border and border costs. Transportation costs are estimated from UNCTAD freight costs for 2019, interview material and documentary research for a typical container shipment worth US\$25,000. Where no data are available, costs are estimated by taking an average from neighbouring countries and applying that to the country concerned. Informal costs were available from documentary research for some African countries; to derive an estimate for all Commonwealth countries, an average across available data for Africa was taken and then weighted up or down using the Transparency International Corruption Index for the specific country. As a result, informal costs shown are relative rather than absolute but serve as a useful benchmark in the absence of other sources. These costs are illustrated in Figure 1.

**How can these costs be reduced across Commonwealth nations?** A base year of 2019 was taken since this was the last full year of data available. Documentary research, semi-structured interviews and the survey of banks and businesses conducted for the ICC were analysed for cost improvements if electronic transferable records were enabled through law. The average across all of these processes produced an estimate for cost reduction that was applied to the data derived from three above and distributed using a net-present value formula over a 5-year period. Cost reduction differentials between nations come from the differential weightings applied to each country reflecting their cost and corruption base.

**How does streamlined access to finance from electronic transferable records impact trade growth?** A baseline forecast of trade was taken by using the Coriolis Technologies refinement and forecasting algorithms (see Technical Appendix). To estimate the effect that cost reductions would have, the following procedure was followed: The total cost per container was taken from adding border, compliance, transport and informal costs together for each country. The average number of containers worth US\$25,000 was calculated by dividing the total trade value for 2019 by the number of containers. The average cost per container was calculated by dividing the average number of containers by the total cost per container. Total trade revenue was assumed to be the total value of trade for any given year from 2019 and projected to 2026. The base projections from the Coriolis Technologies forecasting algorithm were derived from the net present value (NPV) coefficient, which was a geometric average of the total estimated cost reduction spread over the time period AND the cost elasticity of trade for each Commonwealth country taken from the literature search.

## A note of caution

There are two caveats to this research. The first is a bias towards goods trade. This is inevitable given the availability of data, but, as Figure 2 shows, for some SIDS in particular, the dominance of service sector trade means that it is not possible to provide a completely accurate estimate of the improvement in trade or, indeed, the impact on costs except qualitatively through interviews.

Service sector trade data are notoriously incomplete, and this chart contains data for services for the most recent year rather than for 2019 for some countries. Nevertheless, it points to an interesting picture – for some SIDS, notably Vanuatu, Tonga, Samoa, Fiji and Barbados, services (predominantly tourism) are more than 50 per cent of total trade. For other SIDS and LDCs, the low levels of reporting on services trade suggest that there is a lot of cross-border activity that is informal or cash-based and therefore is not appearing in official international statistics.

The second caveat is that all of the estimates are derived from survey-based material or from interviews and documentary research/meta-analysis. There is very little attempt directly in any of the writing about the economic impact of digitisation to provide any values. Indeed, the surveys conducted in this field are estimates of benefits such as reduction in fraud, or streamlined processes, but purposely do not provide quantitative estimates. As a result, the method here represents a first step towards understanding the economic impact but not the final word on the subject.

*Figure 2. Commonwealth reported goods and services exports by country, 2019 (%). Source: Coriolis Technologies.*

## Section 3: Trade across the Commonwealth – the pandemic and its aftermath

---

### Comparison to previous crises

Trade and supply chains across the world have been affected by the Covid-19 pandemic but because of the high dependency on services amongst the Commonwealth's SIDS and the high level of commodity dependency in many African economies, the impact has been felt particularly severely in two ways:

**Reduction in travel and tourism and supply chain disruption:** UNCTAD estimates that travel receipts have dropped by around 70 per cent during the Covid pandemic. As services, and particularly tourism, are said to account for around 25 per cent of SIDS' GDP, and as many of them are highly dependent on imported goods, this makes them especially vulnerable to the key features of the pandemic – a drop in travel and supply chain issues.

**The collapse of commodity prices at the beginning of the pandemic:** this affected inward capital flows, particularly in Africa, and tightened the financial conditions across the continent as the balance of payments deteriorated. This was particularly severe at the start of the pandemic and caused correspondent bank relationships with outside providers of trade finance to fall. Thirty-eight per cent of African banks and 30 per cent of foreign-based banks reported an increase in the numbers of rejections for trade finance from African businesses as a result.

Asian economies have proved more resilient, not least because of the greater focus on manufactured goods in their trade profile.

Goods trade across the Commonwealth fell back in 2020 by around 12 per cent in value terms (Figure 3). What is interesting, however, is that in goods trade terms, this is not as great an impact as the Global Financial crisis when exports fell by 23 per cent and imports by 19 per cent or the oil price collapse between 2014 and 2016 when trade fell by 15 per cent for exports and from and to Commonwealth countries fell back by 22 and 12 per cent for imports. The drop in 2020 was similar: 12.2 per cent for exports and 12.7 per cent for imports.

The effect of course is that for the whole of the Commonwealth, but particularly for import-dependent of commodity trade dependent LDCs and SIDS, the effect has been a worsening trade balance and external debt position, and has highlighted their vulnerability to external shocks. As a result the risk appetite for banks to provide trade finance to exporters in these nations has deteriorated and while Multilateral banks and Export Credit Agencies are now working to address the issues of the pandemic through guarantees to working capital loans and export finance generally, there is a risk that the damage will have longer-term consequences.

*Figure 4 shows the compound annualised growth between 2015 and 2020 and the projected growth in goods trade between 2021 and 2026 by the Commonwealth nation.*

What is remarkable from this is just how many economies, irrespective of the state of development are likely to struggle to recover quickly. Trade across the whole of the Commonwealth is likely to grow in 2026 by just 0.2 per cent annually over the next 5 years, which is modest, but of the 20 countries with the slowest growth projections for 2026, only seven are not SIDS and six of these are emerging African economies.

This is just the trade for goods and excludes the services trade of these nations, but this serves to highlight a point. Trade finance is predominantly focused on goods rather than services trade for the simple reason that it is easier to identify, insure and monitor a goods transaction than a service one. As a result, it is those economies with a strong goods trade presence across the Commonwealth who will be most affected immediately by the pandemic; those economies with a greater exposure to the informal and services economies will have been affected by the reduction in risk appetite generally alongside the drop in travel and tourism.

*Figure 3. Commonwealth total trade, 2006–2020 (US\$ billion). Source: Coriolis Technologies. Figure 4. Commonwealth actual and projected growth in exports by country, 2015–2020 and 2021–26 (CAGR %). Source: Coriolis Technologies.*

## **Where are the emerging trade challenges across the Commonwealth?**

All of this suggests that there are issues in trade across the Commonwealth economies. Our discourse analysis shows the prevalence of words associated with trade that arise commonly across the Commonwealth in the 500 major media sources and publications that were scraped (Figure 5). The results confirm some of the discussion above:

That over a five and a half-year period, the over-reliance on manual processes and inefficiencies because of incorrect paperwork have been the most important issues reported in trade.

That over half of the mentions were in 2021 (to the end of October, the cut-off point for scraping), for most issues suggesting that all had become more important during the Covid pandemic and its aftermath. Over-reliance on paperless trade and inefficiencies through incorrect paperwork were the two that had grown the most in 2021. Only payment delays and human error were issues where the 2021 value was less than half of the total value from 2015–2021.

That for emerging economies digital trade is not the only issue in trade facilitation – logistics, infrastructure, fraud and cross-border barriers are also important.

More broadly, the discourse analysis also shows an increasing association of the word "trade" with different forms of technology (Figure 6). In other words, across the Commonwealth Smart Automation, Blockchain and Artificial Intelligence are increasingly positively associated with the word "trade". The same pattern is illustrated in Figure 6 as shown in Figure 5: that there is almost as many mentions in 2021 as there have been in the whole of the previous 5 years.

*Figure 5. Emerging issues in trade across Commonwealth (discourse analysis). Source: Discourse analysis of 500 media outlets across the Commonwealth. Issues tracked include over-reliance on paper trade, inefficiency through incorrect paperwork, logistics, infrastructure, payment delays, human error, fraud and cross-border barriers. Figure 6. Trade facilitation through technology (discourse analysis). Source: Discourse analysis of 100 media sources. Technologies tracked include smart automation, blockchain technology, artificial intelligence, electronic signatures, data analytics, electronic payment undertaking and innovative APIs.*

What this tells us is that there are challenges in Commonwealth trade and that these are increasingly identifiable, in the trade data and in the public discourse. The interesting and key conclusion from this, however, is that within the public discourse there is a general acceptance of technological solutions associated with trade. In other words, the public would be amenable to the implementation of digital trade, especially if it helped to address some of the barriers to Commonwealth trade also identified in through discourse analysis.

The interviews we conducted corroborated the issues with Commonwealth LDC and SIDS trade and can be summarised as follows:

Interviewees, especially in the Caribbean and in Asia, were concerned about the increased costs of trade. Imports and supply chain delays were key cost concerns, especially in SIDS that were heavily reliant on imports to continue to service their markets.

Interviewees from LDCs and SIDS reported similar challenges in recovering through trade-led growth: excessive costs for exporters, under-developed digital systems and large numbers of microbusinesses excluded from accessing traditional trade finance.

The interviewees were generally positive about the use of legal reform as a trigger to facilitate trade finance and reduce costs. The general viewpoint was that the costs of infrastructure were very high and almost impossible to implement, but enabling regulatory reform would provide the private sector with an incentive to use new technology to create new markets, including for services. It was suggested that the technology is available to do this without major investment from governments.

Caribbean respondents, in particular, saw the Covid pandemic as a primary catalyst for change. They pointed to initiatives in the Eastern Caribbean to promote a Maritime Single Window to ease border controls and the political will, particularly in Jamaica and Antigua and Barbuda, reap the full advantage of more streamlined trade through this mechanism. This, it was argued, would help exporters to become more "agile" in creating new markets both within the region and outside of it.

Public sentiment across emerging Asia, particularly Bangladesh and Malaysia, was seen as a potential obstacle to rapid adoption. The risks of data sharing were highlighted as important "trust" hindrances for businesses, while in some economies, the costs of moving from the informal to the formal sector through greater digitisation would have to be mitigated through incentives – for example, greater access to export support financially or materially.

The prohibitive costs of trade in Africa were a dominant feature of discussions with a particular focus on pre-border, border and informal costs associated with crossings. Respondents were keen to see the AfCFTA work with economic regions and Heads of State to implement Single Windows, Digital identities and electronic documentation but saw the advantages as being associated with greater opportunities to facilitate inclusive trade generally rather than specifically greater access to trade finance.

The interviews, the discourse analysis and the documentary research all point to the fact that these issues can be addressed through electronic transferable records but that for emerging economies the process is considerably more complex and associated with enabling infrastructures, policies to promote inclusion and equal opportunities, and mechanisms to address physical obstacles to trade as much as with the need to reduce the manual components of trade. However, the key thing was that respondents saw the process of legal reform to enable transferable records as being a necessary, if not sufficient, condition to reach the full potential of Commonwealth trade.

This overview can be summarised into three critical success factors. First, the research suggests an emerging political will in the African, Caribbean and the Asian economies, which is supported by the banking sector. While there are evident challenges ahead, structures like the AfCFTA and the Commonwealth, as well as the ICC, were seen as important drivers of the momentum. In Africa, in particular, respondents felt that it was the political leaders at a country level that should be driving the process of reform, not least because it was an effective and cheap way of stimulating change.

Second, the process of standardisation must begin soon – without common standards such as those proposed by the ICC, for example, the potential represented by legal reform would be hard to achieve. However, and equally important, the Commonwealth is a diverse set of nations and it is vital to point out that although it might be possible to create common standards with careful coordination, it is not a Free Trade Area or a Common Market, and hence taxations and regulatory systems would remain divergent across the grouping.

So, third, what this suggests is that it is not the regulations or the infrastructures as such that will need to be standardised. Rather it suggests that there have to be a common set of trade rules globally and a standard set of approaches to amending the law to ensure that the regulations that would underpin greater use of transferable records could be harmonised.

## Section 4: Cost reductions from the introduction of electronic records

All of the above presents a compelling background case to the process of legal reform and which should be a wake-up call to Heads of State. The quantitative analysis presented in this section addresses the primary issue for exporters from LDCs and SIDS of the current prohibitive costs of cross-border trade. It suggests that around US\$90 billion could be added to Commonwealth trade if border and transportation costs, alongside the costs associated with the informal economy, could be reduced.

What is clear is that costs are prohibitive for countries across the Commonwealth (Figure 7).

What is remarkable about Figure 7 is that there are four economies for whom costs are higher than 100 per cent of the revenues received from trade: Vanuatu, Tonga, Gambia and Papua New Guinea. Another 34 economies have costs that are more than 50 per cent of the revenues they receive. These are just the costs associated with border crossing and transport, so adding in raw materials, production, sales and distribution costs would potentially mean that more countries would find the costs of exporting prohibitive. Since border and transportation cost are the easiest to address through Single Windows and electronic transferable records, this suggests that there would be real advantages from implementing the legal reforms to enable this as soon as possible.

*Figure 7. Trade costs as a share of trade revenues 2019. Source: Author's calculations from Coriolis Technologies trade data, Ease of doing business data and meta-analysis.*

### Border costs

Currently, the irregularities across borders, long waiting times to get through border crossings, and the need to produce appropriate travel and transportation documents make cross-border trade difficult. On average, across the Commonwealth, it is costing around US\$521 for a shipment of US\$25,000 (roughly equivalent to a 40-foot container) (Figure 8).

The implementation of smoother border crossings alone would reduce costs to an estimated US\$99 per US\$25,000 shipment. This is an estimated average 81 per cent reduction in border costs.

### Reducing border compliance costs

The amount of paperwork associated with trading across borders costs on average US\$79 per US\$25,000 shipment. This seems relatively small, but as many businesses engaged with crossing borders may be carrying much smaller shipments this is still a cost that needs to be reduced through Single Windows and Digital Identities. The other advantage of this approach, particularly in Africa, according to interviewees, is that these costs are disproportionately high if the exporter cannot read – something which is especially common amongst women. A Single Window simplifies the process and removes the need for high levels of literacy by allowing pre-border completion and automatic identification (Figure 9).

Interestingly, Singapore's costs are relatively high – which may explain why it has been keen to move towards electronic transferable records and streamline its system so quickly. The calculations suggest a reduction in costs on average of just under 80 per cent from digitising the processes of documentation associated with border crossing, rules of origin and incoterms and according to interviewees would particularly help excluded groups and microbusinesses.

*Figure 8. Per US\$25,000 unit reduction in border costs by 2026 with digital trade facilitation (US\$). Figure 9. Per US\$25,000 unit reduction in documentary compliance costs by 2026 with digital trade facilitation (US\$). Source for both: Author's calculations from Coriolis Technologies trade data, Ease of doing business data and meta-analysis.*

## Transport costs

All interviewees spoke of the challenges that Covid had presented for supply chains, pushing up import and transportation costs. SIDS are particularly vulnerable, but costs have risen across the world. These costs are made worse by the potential for border fraud and corruption, making the costs of transporting goods between countries extremely high. Figure 10 combines these two types of transportation costs to create an average cost across the Commonwealth of US\$16,924 per US\$25,000 shipment.

For many SIDS and Belize, these costs are prohibitive, but again, making processes digital and reducing the likelihood of informal costs, could catalyse the most significant cost reduction in value terms even though the percentage reduction is 75 per cent, which is lower than the reductions in other costs. Especially for smaller traders, this could have the biggest single impact for the simple reason that the costs are so high in the first place.

*Figure 10. Per US\$25,000 unit reduction in transport costs (including informal) costs by 2026 with digital trade facilitation (US\$). Source: Author's calculations from Coriolis Technologies trade data, Ease of doing business data and meta-analysis.*

## Bringing it all together – the benefits for businesses from electronic transferable records

Clearly, it is not the most developed markets, such as the UK, Australia, New Zealand or Canada, who would feel the biggest benefits reduction in costs in trade terms. In these countries, production is at the higher end of the value and supply chain and so trade is less cost elastic meaning that at a country level, the total effect is smaller. Rather, it is in the market creation and enabling impact in LDCs and SIDS where the impact would be most felt (Figure 11).

For a country like Uganda, where costs are high and the reduction one of the largest, the effects could be game changing. One interviewee pointed out the challenges currently faced: that border crossings were impossible without some form of informal payment and that there were no easy means of filling in documentation to ensure that common standards were followed. Uganda appears to be one of the countries that would benefit most from the reduction in transportation costs.

*Figure 11. Reduction in transportation costs as a share of total trade, 2019 and 2026. Source: Author's calculations from Coriolis Technologies trade data, Ease of doing business data and meta-analysis.*

*Figure 12 shows the percentage reduction between 2019 and 2026 in costs as a percentage of total trade. It shows a slightly different take on the same data and suggests that:*

The SIDS are the dominant beneficiaries in terms of the total percentage reduction: seven of the top ten beneficiaries are SIDS.

Belize will benefit – perhaps because it is one of the countries with the highest informal costs associated with trade.

Malaysia is one of the top ten beneficiary nations and as its costs are known to affect its overall levels of trade, this could have a substantial beneficial impact on trade values and volumes as well.

## Summary

As has been stressed throughout, these are stylised effects based on estimates and modelling from various databases. That said, there are some clear messages that come out of this analysis:

The reduction in total transportation costs is over 75 per cent across the Commonwealth, and this in itself will represent huge advantages for exporters.

Addressing complexity in the system through digital IDs and Single Windows will enable excluded groups to access trade more readily.

The reduction in informal costs comes from the paperless trade process and while there may be some short-term issues in building trust to encourage informal actors into the formal trade system, reducing these costs will disproportionately help exporters and potentially improve revenues for governments as well.

*Figure 12. 2020–26 Reduction in trade costs from digitisation as a share of total trade (%). Source: Author's calculations from Coriolis Technologies trade data, Ease of doing business data and meta-analysis.*

## The impact on trade

The research set out to estimate the impact on trade of electronic transferable records. It is important to point out again that the estimates are derived from qualitative assessments and represent a target from potential market creation, the inclusion of more diverse exporters and the elimination or significant reduction in some of the non-trade barriers to trade.

The underlying base forecasts are derived from the methodology described in the Technical Appendix. Estimates for trade increases over a 5-year period from interviews, meta-analysis and the bank survey have been averaged and the coefficients which are weighted for the elasticity of trade by country and allocated on the basis of an exponential formula that assumes relatively slow take up in the first 2 years.

The effects are nevertheless striking and illustrated in Figures 13–18.

*Figure 13 shows that, by combining the impact of cost reductions and greater access to finance, there could be as much as a US\$1.1 trillion increase in trade by 2026 assuming the upper bound of estimates by 2026.*

The individual country results are broken down by region and illustrated in Figures 14–18. The current trade and the upper and lower bound estimates are given for each. The larger economies are removed from each regional grouping for illustrative purposes since they distort the numbers for small economies.

*Figure 13. How trade would accelerate through electronic transferrable instruments (e-bills, promissory notes): trend and forecast growth from digitisation compared (US\$ billion). Figure 14. Africa (excluding South Africa). Figure 15. Americas and the Caribbean (excluding Canada). Figure 16. Asia (excluding India and Singapore). Figure 17. Pacific (excluding Australia and New Zealand). Figure 18. Larger economies. All show the boost to trade from digital trade facilitation at borders (trend, best and worst case by 2026). Source: Author estimations based on Coriolis Technologies machine-learning-based forecasts and interviews, meta-analysis and survey data.*

## Summary

The results show clearly that for all economies there will be a significant improvement in trade resulting from two things:

the reduction in costs, which enables more exporters to access trade routes;

the improvement in access to finance, which has the effect of creating markets, especially for SMEs who are currently excluded because they are unable to access traditional forms of trade finance because of the due diligence costs involved.

## Section 5: Summary and Concluding remarks

This research has covered 54 diverse economies and presented a picture of the potential for cost reductions and trade increases as a result of the introduction of legal reform to enable the use of electronic transferable

records across the Commonwealth. It has shown that costs will, on average, fall by around 75 per cent and for some costs as much as 81 per cent. This itself could enable as much as US\$90 billion in additional trade but, if combined with measures to use electronic transferable records in trade finance, will have a multiplier effect enabling a total of nearly US\$1.2 trillion across the Commonwealth. In some countries, such as Vanuatu, where the costs of trade are disproportionately high relative to the revenues received, this has the effect of creating trade where previously very little existed.

As has been highlighted above, there are limitations to this research, which come from the fact that there are very few precedents of this type of research and therefore benchmark figures have had to come largely from qualitative sources. Nevertheless, the size of the trade cost reductions and the market creation effects should act as a Call to Action for policy makers across the Commonwealth, especially if supported by other policies to promote education, standardisation and trust-building between the Commonwealth itself, the International Chamber of Commerce, the Digital Standards Initiative, the International Trade and Forfeiting Association and the African Continent Free Trade Area. This will enable standards and implementation to accelerate and the benefits highlighted here to be accelerated.

## Technical Appendix: Coriolis Forecasting Methodology

---

Coriolis trade flow refinement methodology is a set of algorithms to provide current and accurate global trade flow estimates. The aim is to provide a comprehensive solution to two known error sources in trade data – data lags and bilateral asymmetries. Lags arise from the complexities of reporting and aggregating trade for all goods in a country. The fastest reporting economies are still at least 3 months lagged. Bilateral asymmetries in trade data (both in goods and services) are a well-known phenomenon in official statistics. They occur when the reported exports from country A to country B do not match the reported imports to country B from country A.

We define the two challenges as statistical and machine learning problems. We tackle lags through a forecasting model based on Recurrent LightGBM. This step solely calculates the current and future (12 months ahead) monthly reported trade values (USD) of all global trade flows on a 6-digit HS-code level. Modelling reported trade values directly means that the model forecasts "what the Reporting economy will report". If the economy constantly under-reports, then the model will also under-report in the same manner. The reporting bias is estimated and mitigated in the subsequent steps: trade value refinement. In this part, we design a Bayesian methodology approach to quantify reporters' trustworthiness. The trustworthiness score is then used to determine the actual value of trade between two economies by weighting the reported export value and the reported import value based on their score.

Lastly, there is also the challenge of scale. There are 53 million trade flows, which require a monthly update. This has implications for the modelling logic and for the choice of hardware and data-processing methodologies. The forecasting model must be able to create robust forecasts for trade flows of all-time series types, including very sparse trade flows, smooth trade flows, jumpy, erratic or intermittent trade flows without the possibility of making model changes for a single trade flow.

The base forecasting below is used to create a view of what will happen to trade if all else remains equal. There are no assumptions, but the machine learns for the history of a specific time series.

The diagram summarises the flow of data in the Coriolis trade flow refinement approach. On top is the raw Comtrade data. Two example trade flows are defined:  $\text{Export}(C | A \rightarrow B)$  the reported export value of commodity C from Reporter A to Partner B and  $\text{Import}(C | B \leftarrow A)$  the reported import value of commodity C by Reporter B from Partner A. The pair of those values is called a symmetric trade pair. In the long run (summed over many years), these values should be the same because both relate to the same commodity just

registered at different times and in different countries. In practice, there are differences in reporting, goods are qualified differently, pass through third countries or different exchange rates are used to qualify the value. In the medium and short run, there are also differences stemming from shipment delay, warehousing and the treatment of differential reporting – Freight on Board (FoB) for exports and Cost Insurance and Freight (CiF) for imports estimated as up to 5 per cent of the asymmetry in any one trade flow (Baranga 2018).

The second row represents the modelling phase. From the raw data, two models are built. One is the Bayesian trustworthiness model, which outputs a trustworthiness score for each Reporter and Commodity combination. The second is the forecasting model, which uses all previously available trade flow data, additional features derived from the raw data and external variables (Bill of Lading, PMI, macroeconomic data) to estimate current (nowcast) and future (forecast) values of the raw trade flow data.

The third step is to put together the two models to arrive at the actual trade values for the past, present and future trade.

## Forecasting methodology

*Karol Przybylak and Jintao Long, Coriolis Technologies*

The goal is to create a forecasting algorithm for monthly bilateral trade flows at a 6-digit HS code level (note that this can then be aggregated to four-digit and two-digit levels and from monthly data to annual data). Trade flow data are registered by the customs offices of each country and are then processed and aggregated by UN Comtrade as the most comprehensive and consistent "raw" data source for world trade.

The Comtrade data are lagged by several months; hence, the goal is to create "nowcasts" – estimates of the current values and forecasts – estimates for 12 months ahead on the moment of forecast creation. The algorithm is based on Machine Learning and Statistical models for time series. It extracts patterns in trade flow time series and patterns between the trade flow time series and external variables (macroeconomic indicators, trade restrictions, etc.) to extrapolate those patterns into the future and hence create forecasts.

## Setting and data

See Table 1. The most prevalent characteristics of the task are:

The data are big – 58 million trade flows with monthly data and hence 58 million forecasts to be updated each month.

The data are lagged – even creating "nowcasts" means forecasting ~6 months ahead. A lot can happen in trade within 6 months.

There are different sources for the reporting lag and the lag can also be changed dynamically.

Exogenous variables have different lags. Some, like for example Bill of Lading data, can be used as a proxy for trade flows in the more recent periods where trade data is not yet available.

The data are hierarchical. Commodities are categorised into six-digit HS codes. This hierarchy can help utilising cross-correlations between commodities with similar codes.

There are trade flows that look "censored" – with periods of 0s and then some months with data. The algorithm has to be able to output positive or 0 numbers only. In machine learning terms, it is a truncated regression problem.

*Table 1. Summarises the input and output variables of the model. Light blue cells (T-7 to T-1) are historic values available at time point T before creating the forecast; light green cells (T to T+6) are values to be forecasted by the model. Vbaec, the external variable, is a total monthly USD value of the export e of commodity C between reporter country B and partner country A.*

| Variable / Time | T-7  | T-6  | T-5 | T-4 | T-3 | T-2 | T-1 | T | T+1 | ...T+6 |
|-----------------|------|------|-----|-----|-----|-----|-----|---|-----|--------|
| Vbaec           | 940k | 980k | ?   | ?   | ?   | ?   | ?   | ? | ?   | ?      |

| Variable / Time                        | T-7  | T-6  | T-5  | T-4  | T-3  | T-2  | T-1  | T    | T+1  | ...T+6 |
|----------------------------------------|------|------|------|------|------|------|------|------|------|--------|
| GDP forecast or other macro indicators | 10bn | 10bn | 10bn | 10bn | 10bn | 10bn | 10bn | 10bn | 10bn | 10bn   |
| Bill of lading data                    | 1M   | 1.1M | 1M   | 0.9M | 1M   | 1.2M | 1.1M | -    | -    | -      |
| Oil prices                             | 40\$ | 41\$ | 46\$ | 34\$ | 42\$ | 37\$ | 38\$ | 41\$ | -    | -      |
| Trade restrictions and tariffs         | 20%  | 20%  | 20%  | 20%  | 20%  | 20%  | 20%  | 20%  | 20%  | 25%    |

## Literature and Resources

This section summarises the available research in the context of trade flow forecasting. For a full list of resources, including research papers and publically available code, refer to the Bibliography section.

Research from two main areas of interest has been revised: first, Machine Learning and Statistical approaches to trade flow forecasting. This helped gaining a domain-specific understanding of the problem. Second, we analysed research on large-scale time series forecasting, which could be applied to any time-series data.

### Machine Learning and Statistical approach to trade flow forecasting

In the first instance, the sources focus either on explanation of the model (parameter analysis) or on forecasting accuracy. Traditionally statistical and econometric literature is focused on explanation and the significance of results while the Machine Learning community is focused on accuracy. We are interested in the later but the results of statistical research provide great insight into the data, which help to understand the problem and propose improvements also for the accuracy focused approaches. None of the research so far has focused on forecasting all trade flows globally. So far, it has always been a small selection of commodities and a small selection of countries.

The best-performing models from the reviewed research usually share the following characteristics:

- ability to extract nonlinearity;
- ability to extract patterns in cross-correlations between different trade flows;
- ability to extract seasonality and change points;
- usage of external regressors;
- handling of a dynamically changing time structure;
- handling very different types of time series. Some trade flows are 90 per cent 0s;
- handling large amounts of trade flows (58M) simultaneously;
- handle relatively small amount of data for specific trade flows. In the best case, we have 140 months of data for a given trade flow. Overfitting is a real possibility. The model should operate on a larger set of time series with shared parameters.

Deep Learning and/or on Boosting algorithms (Random Forests, XGB, Lightboost) are pretty much the only available approaches to handle all of these issues.

### Large-Scale Time Series Forecasting – Model Estimation

LightGBM is chosen as the supervised learning model for forecasting. The model is built together for a set of trade flows. This is a necessity as it would not be possible to be built at once for all trade flows together with the currently available infrastructure. An optimisation for the choice of the grouping remains to be done. Generally, an informed clustering should be better than either of the extremes – one model for all trade flows or one model for each trade flow.

This implies that a group-invariant error measure has to be chosen as the algorithms' loss functions. With standard error measures like AE, RMSE trade flows with higher values would have a stronger impact on the

error measure. Hence, the learning algorithm would optimise the parameters towards mostly performance on those flows. MAPE is group-invariant but not scale-invariant. That is, if the actual value is close to 0, then even if the forecast is close in nominal terms in percentage terms it could be millions per cent off the actual value. Again, this outlier would have unproportionately large impact on architecture of the LightGBM model. Currently, Mean Absolute Scaled Error is chosen as a measure that is both scale-invariant and group-invariant.

LightGBM, as a tree-based ensemble algorithm, does not extrapolate into the future. Unlike generative models, the values for different time horizons cannot be generated by one model. We built one model for each forecasting horizon:  $[M = f(X_t, Y_t + 1), M = f(X_t, Y_t + 2), \dots, M = f(X_t, Y_t + s)]$ , where  $s$  is the required horizon.

To capture recurrency and other time-series related characteristics, past values of the trade flow are fed into the model as exogenous variables. The set of exogenous variables  $X_t$  consists of past values of the modelled trade flow, past values of other trade flows (including those of the symmetric trade pair), month (1–12), period (201001-max), Bill of Lading values of the related HS code, macroeconomic indicators, tariff values, non-tariff related trade measures for the related time-space, time-series characterisation features (intermittency, smoothness, etc.), macroeconomic indicators and USD exchange rate-related features. Since multiple trade flows are fed into one model cross-correlations between reporters, partners, commodities and trade direction (export/import) are latently inferred by the model without explicitly defining them as exogenous variables.

The model is estimated on data from January 2010 to December 2017 and tested on all months of 2018. Hyperparameter optimisation techniques are utilised to tune the performance, and then the model is re-estimated on all available data from January 2010 – now and forecasts are created for all months until May 2022.

### **Missing and zero trade values (especially significant in the context of emerging markets)**

The available trade data have a fraction of 0 or close to 0 monthly trade values. It also has a bigger fraction of months where certain trade flows were not reported. In this situation, the trade might not be reported properly, it might be reported with the next month data or it might be a custom of the reporter to not report values if they are 0. The forecasting algorithm is reused to also decide on the missing values whether to treat them as 0s or a missing value and estimate that value.

The algorithm is also informed with results of an additional investigation on the randomness of missing values. Missing value in global trade data is used as an indicator of the level of pattern intervention presented in a particular trade series. Three categories of data missingness are developed – missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR). In a time series, the missing value proportion and the maximum missing gap by month are integrated into a system to determine missingness types. In general, it is assumed that time series with very low missing value proportion would fall into the category of MCAR, while those with relatively high missing value proportion and maximum missing gap by month would fall into the category of MNAR, and those outside the above two categories would fall into the category of MAR.

### **Bayesian Trustworthiness (TW)**

This is a proposal of a probability model to evaluate the trustworthiness of global trade value on the granular level of reporter and commodity.

Definitions and conditions:

Chg is the acceptable upper limit of a gap difference in percentage, default set to 5 per cent.

A symmetric trade value pair matches when the proportion of gap (PoG) is smaller than the acceptable upper limit (Chg) of the sum of the doubleton.  $PoG = |V_a - V_b| / (V_a + V_b)$ .

An observation (O) is to see whether a symmetric trade value pair matches or not, 1 means match ( $PoG \leq Chg$ ) and 0 means mismatch ( $PoG > Chg$ ).

The relationship between an observation result and country trustworthiness – in an observation  $Ok$ , where country A reports import from country B on commodity C at time T ( $V_{abict}$ ), the symmetrical trade should be country B reports export to country A on commodity C at time T ( $V_{baect}$ ): when country A is trustworthy thus reporting the true value of  $V_{abict}$ , and country B is also trustworthy thus reporting the true value of  $V_{baect}$ , then  $Ok = 1$  (values match). When country A is not trustworthy thus reporting the false value of  $V_{abict}$  and country B is trustworthy thus reporting the true value of  $V_{baect}$ , then  $Ok = 0$  (values do not match). When country A is trustworthy thus reporting the true value of  $V_{abict}$  and country B is not trustworthy thus reporting the false value of  $V_{baect}$ , then  $Ok = 0$  (values do not match). When country A is not trustworthy thus reporting the false value of  $V_{abict}$  and country B is not trustworthy thus reporting the false value of  $V_{baect}$ , then: (d1) the probability of  $Ok = 0$  (not match) is 95 per cent; (d2) the probability of  $Ok = 1$  (still match) is 5 per cent.

Each observation is independent.

Caps are applied with an upper limit of trustworthiness be 99 per cent and a lower limit be 1 per cent. Posterior probability outside this range would be capped.

### Bayesian Steps

The end goal is to calculate the probabilities of every country being trustworthy in reporting trade values for every sector, import and export, based on reported raw trade values.

With Bayesian probability as the ground theory, suppose we want to know all of the probabilities of countries being trustworthy in reporting trade values regarding commodity C. Figure 1 shows all the trade values reported at time T. There are 12 pairs of symmetric trades; therefore, 12 observations can be extracted from this single time point.

Suppose we are looking to calculate the probability of country A being trustworthy in reporting its import trade values of commodity C. According to the Bayesian theorem, we are looking to calculate the posterior probability of country A being trustworthy in import  $P(A_i\text{--trustworthy} | ON)$  given ON, where B represents the partner country in the observation.

$P(A_i\text{--trustworthy})$  is the prior probability, which represents the pre-knowledge before each observation. It is the prior probability of country A being trustworthy in reporting import trade on commodity C. It is equal to a preassigned value in the first iteration of calculation, and generally equal to the posterior probability  $P(A_i\text{--trustworthy} | ON)$  calculated in the last observation.

$P(ON)$  is the marginal probability, which means the probability of ON happening. When  $ON = 1$ , to make  $ON = 1$  happen, either they both are telling the truth, or both lying while values still are matching, the probability of which is set to 5 per cent as defined in condition 4d. When  $ON = 0$ ,  $P(ON = 0) = 1 - P(ON = 1)$ .

### Algorithm Pseudo Code

Therefore, the final algorithm is as follows:

For commodity S1 to SN: For time point/month T1 to TM: Pre-assign prior probability to every country (global default 0.5, half trustworthy); Repeat: For country C1 to CL: Calculate  $P(A_i|e, \text{trustworthy} | ON)$  with prior probability; Update prior probability with  $P(A_i|e, \text{trustworthy} | ON)$ ; Until all the probability converges without dramatic changes; Average all probabilities for each country through different time points (or make them all converge to one);

## Refinement – trade report latency calculation

We define latency as the lag between reported export and import of the same trade flow. It is an integer value describing the number of months it takes for goods to be shipped between the countries. In reality, this number will change dynamically.

**Approach 1:** OLS approach to investigate latency for each reporter/partner/commodity. OLS to calculate correlation between shifted series ( $T + N$ ,  $N = 1, 2, 3, 4, 5$ ). Choose  $N$  with lowest p-value. When  $>1$   $N$  are found to be significant, choose largest  $\text{abs}(t\text{-value})$ .

**Approach 2:** Allocated lagging level to be 1, 2, 3. Assign lagging level to reporter/partner pair based on geographical coordinates.

**Step 1.** For each set of reporter country  $C1$ , partner country  $C2$ , commodity  $S$  and period  $T$ , identify the genuine import and export pair of trade records by applying the pre-calculated trade latency  $N$ , such that:  $X1$  = the import of  $S$  from country  $C1$  from country  $C2$  at period  $T + N$ ;  $X2$  = the export of  $S$  from country  $C2$  to country  $C1$  at period  $T$ .

**Step 1 (version 2).** In this approach, we intend to incorporate the fact that commodities arrive from exporter to importer with a delay. For example, let's assume that Japan reported exports of TVs to Belgium in March to be 1000USD. This is a total of all shipments registered by the customs office. Also, they could have been registered anywhere between 1 and 31 March. The shipments could be exported by plane or by ships and from all different ports. The ships have different speed, and shippers use different routes. All this leads to the shipments arriving in Belgium at different times. Some will arrive in the same month and some will arrive months later.

Using the same definitions as above:  $X1(T)$  = the import of  $S$  from country  $C1$  from country  $C2$  at period  $T$ ;  $X2(T)$  = the export of  $S$  from country  $C2$  to country  $C1$  at period  $T$ . Through a modelling process, we seek to substitute  $X2(T)$  with an optimised  $X2^*(T)$  such that  $X2^*(T)$  is the weighted average of export of  $S$  from country  $C2$  to country  $C1$  at period  $T, T-1, T-2, \dots$

The problem remains to estimate the optimal weights. It is helpful to think of this as a linear regression (equation 1):  $X1(T) = w_0 + w_1 \cdot X2^*(T) + w_2 \cdot X2^*(T-1) + w_3 \cdot X2^*(T-2) + \dots$

It is tempting to define  $X2^*(T)$  as just the RHS of this equation. Although this has two major drawbacks. First is that this would mean that we modelled import values with export values. In other words, through a linear model, we made the import values as similar to the export values as possible. But this is not really our goal because in this step we just want to solve the lag structure. The problem of defining where the actual value lies between  $X1$  and  $X2$  is solved in Step 2. Second, with OLS, the weights may achieve negative values. We would like to impose that a reported export will always cause an increase in the reported import (somewhere in time). We can also impose  $w_0 = 0$  because we know that if all  $X2(T) = 0$  at all  $T$ , then  $X1(T)$  also has to be 0. In other words, if there was nothing exported, import should also be 0.

Taking these into account, let's define (equation 2):  $X1(T) = w_1 \cdot X2^*(T) + w_2 \cdot X2^*(T-1) + w_3 \cdot X2^*(T-2) + \dots$ , where  $w_i > 0$ . The weights can be estimated through non-negative least squares, and there is also a corresponding Python implementation. If we sum the weights  $w$ , then we achieve a general over/under-reportedness measure of  $X2$  compared to  $X1$ . That is if  $\|w\| < 1$ , then  $X2$  overreports compared to  $X1$ , otherwise it under-reports. Now we do not want to get involved in over/under-reportedness in this step, so above we should normalise  $w_i$  to sum to 1. Now substitute  $X2$  with  $X2^*$  and continue with Step 2.

**Step 2.** For each pair of trade, the steps are followed through. Calculate the percentage difference between the different reported values of the same trade flow as:  $d = |X1 - X2| / (X1 + X2)$ . Based on the scale of  $d$ , apply the following fixing:

If  $X_1 = X_2$  ( $d = 0$ ), no change is made.

If  $d > 0.25$ , the difference is large so that we choose the number reported by country with larger trustworthiness index for this commodity and trade flow.

If  $0 < d < 0.25$ , the difference is mild so that we set the refined value to be the weighted average by trustworthiness:  $(TW_1 \cdot X_1 + TW_2 \cdot X_2) / (TW_1 + TW_2)$ .