Does Digital Trade Change the Purpose of a Trade Agreement?*

Robert W. Staiger
Dartmouth and NBER

December 9, 2021

Abstract

The design of a trade agreement should reflect its purpose. Does digital trade change the purpose of a trade agreement? To explore this question, I first describe the definitional and classification issues associated with digital trade, and for modeling purposes I adopt a simple taxonomy of the ways in which digital trade can arise and the policies that can be used to restrict such trade. I then review what the theoretical literature on the economics of trade agreements has to say about the purpose of a trade agreement in a pre-digital model world economy, and how this purpose can be seen to be reflected in the broad design features of both GATT and GATS, the WTO agreements that govern international trade in goods and services respectively. Finally, I introduce digital trade into the model world economy and revisit the purpose of a trade agreement. From this perspective I consider whether the rise of digital trade warrants changes in the design of the WTO.

*This paper was written to serve as a background paper for the Roundtable "The Role of Trade Agreements in the Regulation of Trade in Data," at the Rotman School of Management of the University of Toronto on March 26 2021. I thank Alan Sykes for many helpful discussions and comments, Paul Hager for outstanding research assistance, and my discussants Avi Goldfarb and Dan Trefler as well as seminar participants at the Roundtable for many useful comments. A condensed version of this paper will appear as Chapter 9 in Staiger (forthcoming).

1. Introduction

We live in an increasingly digital world. The Internet and digitalization are fundamentally changing the way people, firms and governments interact, both within and across national borders. Though definitions vary and data is incomplete, economists concur that digital trade has revolutionized the global economy and will continue to do so for the foreseeable future. In 2018, digitally deliverable services comprised 50 percent of overall global services exports, and estimates place the digital share of global GDP between five and 16 percent (UNCTAD, 2019). The consequences and opportunities posed by the rapid expansion of digital commerce are felt by individual consumers, by small firms as well as large ones, and by economic superpowers as well as emerging and developing economies (see, e.g., Castro and McQuinn, 2015).

In this paper I consider the implications of digital trade for the design of the World Trade Organization (WTO), an institution whose main features were negotiated and implemented while the Internet was still in its infancy.[1] Does the importance of digital trade today imply that the approach taken by the WTO to global trade rules is fundamentally out of date?

Certainly digital trade has made possible new forms of trade protection. The US International Trade Commission (USITC, 2014) offers a partial inventory:

> Localization requirements, market access limits, data privacy and protection requirements, intellectual property rights infringement, uncertain legal liability rules, censorship, and customs measures in other countries all present obstacles to international digital trade. (p 14)

And in light of these novel forms of trade protection and the ubiquitous nature of digitalization in the global economy, a recent OECD Trade Policy paper argues that countries need to take a more "holistic" approach to market openness as a result of the rise of digital trade, stating:[2]

> Today, a simple digital trade transaction rests on a series of trade-related factors that enable or support the transaction. For instance, the ability to

[1] This is not to say that WTO members have only recently become aware of the disruptive potential of the Internet for global trade rules. Soon after the WTO's creation in 1995, the member governments agreed to convene a "Work Programme on Electronic Commerce" (WTO 1998), and an initial study of electronic commerce and the role of the WTO was issued in 1998 (Bacchetta, et al, 1998). See WTO (2018) for an updated WTO report on digital trade.

[2] This argument is made widely. See, for example, Ahmed (2019) and Ciuriak (2019).

> order an e-book depends on access to a retailer's website. This in turn depends on the regulatory environment which determines the conditions under which the retailer can establish the webpage as well as on the cost for the consumer of accessing the Internet – a cost which, in turn, is affected by the regulatory environment in the telecommunications sector. The purchase of the e-book will also be affected by other factors, such as the ability to pay electronically and the tariff and non-tariff barriers faced by the physical device used to read the e-book.
>
> A barrier on one of these linked transactions will affect the need or the ability to undertake the other transactions. This means that market openness needs to be approached more holistically, taking into consideration the full range of measures that affect the ability to undertake any particular transaction. For instance, Internet access may be a necessary but not sufficient condition for digitally enabled trade in goods to flourish. If logistics services in the receiving (or delivering) country are costly due to service trade restrictions, or if goods are held up at the border by cumbersome procedures, then the benefits of the digital transformation may not materialise. (Casalini et al, 2019, p. 5).

The WTO has been slow to take up this challenge, and as a result many WTO members have moved away from multilateral efforts to update trade rules for the digital era, and are pursuing these goals in "deep integration" regional and mega-regional agreements instead (Wu, 2017). Whether these agreements can or should be seen as a model for the WTO's approach to digital trade is an open question.

In short, with digitalization permeating so deeply into modern life, it is tempting to conclude that the world is now truly "flat," that everything has changed.[3] But for the specific task of designing an effective trade agreement, can we be sure that everything is now different?

To evaluate the need for a re-design of the WTO in response to the new realities of the digital world economy, I adopt a basic premise from the literature on the economics of trade agreements: namely, that the design of a trade agreement should reflect its purpose, where by purpose I mean the "problem" that the agreement is supposed to "solve" so that countries can eliminate international policy inefficiencies and enjoy the mutual benefits from implementation of internation-

[3] Friedman's (2005) early pronouncement of the death of distance and that the world is now flat gave rise to critiques by economists (see e.g., Leamer, 2007) arguing that in fact distance was alive and well in the trade data (see, e.g., Disdier and Head, 2008).

ally efficient policies.[4] From this perspective, I ask: Does digital trade change the purpose of a trade agreement? The answer to this question can illuminate the nature of the challenge that digital trade poses for the WTO and the world trading system.

I first present a simple partial equilibrium model of trade between two countries in a pre-digital world. I review what the theoretical literature on the economics of trade agreements has to say about the purpose of a trade agreement in this setting, considering both trade in goods and trade in services; and I describe how this purpose can be seen to be reflected in the broad design features of both the General Agreement on Tariffs and Trade (GATT) and the General Agreement on Trade in Services (GATS), the WTO agreements that govern international trade in goods and services respectively. I then introduce digital trade into the model world economy and revisit the purpose of a trade agreement, to investigate whether the problem for the agreement to solve has changed. It is from this perspective that I evaluate whether the rise of digital trade warrants changes in the design of the WTO.

For my analysis, I follow the WTO and define digital trade as "the production, distribution, marketing, sale or delivery of goods and services by electronic means."[5] A key question is how to parsimoniously introduce digital trade into a model world economy. At its core, digitalization impacts the economy through its impact on transaction costs. Surveying the literature on digital economics, Goldfarb and Tucker (2019) observe:

> Understanding the effects of digital technology does not require fundamentally new economic theory. However, it requires a different emphasis. Studying digital economics starts with the question of "what is different?" What is easier to do when information is represented by bits rather than atoms? Digital technology often means that costs may no longer constrain economic actions. Therefore, digital economics explores how standard economic models change as certain costs fall substantially and perhaps approach zero. (p 3)

Goldfarb and Tucker emphasize five kinds of costs that can be substantially lowered in the presence of digitalization: (i) search costs; (ii) replication costs; (iii)

[4] For a recent review of the literature based on this premise, see Bagwell and Staiger (2016).

[5] There are a variety of working definitions of digital trade that have been proposed in the literature (see Meltzer, 2019, for a recent discussion), but for my purposes the WTO definition seems appropriate.

transportation costs; (iv) tracking costs; and (v) verification costs. They continue:

> Search costs are lower in digital environments, enlarging the potential scope and quality of search. Digital goods can be replicated at zero cost, meaning they are often non-rival. The role of geographic distance changes as the cost of transportation for digital goods and information is approximately zero. Digital technologies make it easy to track any one individual's behavior. Last, digital verification can make it easier to certify the reputation and trustworthiness of any one individual, firm, or organization in the digital economy. Each of these cost changes draws on a different set of well-established economic models, primarily search, non-rival goods, transportation cost, price discrimination, and reputation models. (pp 3-4)

A thorough analysis of the impacts of digital trade on the purpose of trade agreements would build from this taxonomy of cost implications, and would consider those impacts in each of a set of economic models that could appropriately micro-found the implications of reductions in each kind of cost.[6]

Rather than working with a modeling framework that captures the micro-foundations of how digitalization can impact any of these particular costs, I take a reduced-form approach and simply assume that digitalization reduces the costs of international trade, and that in choosing their digital trade policies governments weigh the trade-cost reducing impacts that a more open digital environment may bring against any possible non-pecuniary externalities associated with issues of privacy, national security, law enforcement and the like that may accompany such a policy. Admittedly, this approach misses many of the issues surveyed by Goldfarb and Tucker (2019), especially as these issues relate to the impact of digitalization on replication costs, tracking costs and verification costs. My intent is to capture important impacts of digitalization on transportation costs and search costs with this modeling approach, and leave other dimensions of the impacts of digitalization on trade to future research.

My approach to capturing the impacts of digitization on trade is similar to that of Freund and Weinhold (2004), who model the impact of the Internet on trade in goods by assuming that it reduces the fixed costs of entering a particular market. It is also in the spirit of the trade literature that interprets the impact that distance has on trade as reflecting search and information frictions (see, for

[6] See McCalman (2021) for an early paper that develops an explicit model of search and social media to consider the role of trade agreements in a digital world.

example, Allen, 2014, and Head and Mayer, 2014), and that views such frictions as diminished by the Internet and online markets (see, for example, Lendle et al, 2016, for trade in goods, and Blum and Goldfarb, 2006, for trade in services).[7]

My findings suggest that the WTO is perhaps better designed to deal with digital trade than is commonly believed. I show that where the non-pecuniary externalities associated with digital openness are purely local, the purpose of a trade agreement for both goods trade and trade in services is unchanged by the advent of the digital world. As I demonstrate, this implies that the existing "shallow-integration" features of GATT can in principle be applied to digital policies impacting goods trade in such a world; and while GATS is currently structured as a deep-integration agreement, a feature that Staiger and Sykes (2021) argue can account for its relative lack of success as compared to GATT, I show that a GATT-like shallow-integration approach to trade in services along the lines proposed by Staiger and Sykes could in principle be applied to digital policies impacting services trade as well. I also show that where the non-pecuniary externalities associated with digital openness cross international borders, the purpose of a trade agreement is more complex. But I argue that even in this case there may be an approach to integration for goods and services trade in a digital world that lies somewhere between the WTO's shallow integration approach and a fully deep approach where countries negotiate directly over all aspects of digital policy.

The rest of the paper proceeds as follows. Section 2 covers definitional issues associated with digital trade and proposes a taxonomy that I use in the analytical sections of the paper. Section 3 describes the novel features of the policy environment that arise with digital trade. Sections 4 and 5 present the analytical results regarding the impacts of digital trade on the purpose of a trade agreement for trade in goods and trade in services, respectively, while section 6 considers the implications of these results for the design of the WTO. Section 7 concludes.

[7] Blum and Goldfarb (2006) distinguish between "taste-dependent" and "non-taste-dependent" services, and show that distance does not reduce demand for non-taste-dependent services that are web-based, consistent with the notion that digitalization has reduced distance-related trade costs for such services to zero. Blum and Goldfarb also find that taste-dependent services continue to exhibit falling demand with distant even when they are web-based, and interpret this finding as reflecting regional differences in taste.

2. What is Digital Trade?

As mentioned in the Introduction, while a variety of definitions of digital trade have been proposed in the literature, for the analysis below I follow the WTO and define digital trade to mean "the production, distribution, marketing, sale or delivery of goods and services by electronic means."[8] Still, the potential challenges that digital trade poses for the WTO and the world trading system are well-illustrated by considering the struggle to define and classify digital trade. For example, in describing the electronic commerce that comprises the WTO's definition of digital trade, Bacchetta et al (1998) distinguish among three stages in electronic transactions, any of which may occur on an arms-length or within-firm basis: the searching stage, the ordering and payment stage, and the delivery stage. Bacchetta et al elaborate on these stages as follows:

> The searching stage is where suppliers and consumers interact in the first instance. This stage may or may not lead to an actual transaction. The second stage entails ordering and payment for the good or service, typically through the electronic transmittal of credit card or bank account information. The third stage is delivery. Only those transactions that can be concluded through electronic delivery of digitalized information may be carried through entirely on the Internet. Electronic commerce via the Internet must end at the second stage for purchases which cannot be delivered electronically, including physical goods like flowers or bicycles, and services that can only be supplied if the supplier and consumer are in physical proximity, like haircuts, tourism and construction. It is the expanded scope for the third stage of electronic commerce transactions — that of taking electronic delivery of the purchase — which is perhaps the most notable contribution of Internet technology and the most challenging aspect from a policy perspective. (p 1)

Meltzer (2013) notes the breadth of impact on trade that is implied by the three stages of electronic transactions identified by Bacchetta et al:

[8] Throughout this paper I use the term "digital trade" in the same way that the WTO uses the term "electronic commerce," or "e-commerce." The definition of digital trade that I quote in the text is the definition of e-commerce adopted by the WTO's Work Programme on Electronic Commerce. This definition can be found on the WTO webpage at https://www.wto.org/english/thewto_e/minist_e/mc11_e/briefing_notes_e/bfecom_e.htm.

> As these three stages demonstrate, the Internet allows for international trade in electronic goods and services and cross-border data flows also have important indirect effects on international trade. For instance, advertising on search engines such as Google and Bing bring together overseas buyers and sellers and is often how consumers learn of the goods and services available in other countries. Advertising is therefore often a necessary precursor to the online transaction that leads to international trade. The ability for researchers in different countries to share data and collaborate can determine whether an international services trade occurs. The Internet and cross-border flow of data is also crucial for other services that support and enable international trade, such as VoIP - internet based communications through sites such as Skype and email. Cross-border data flows are also necessary for the financial transfer to complete the transaction. (p 11)

Classifying digital trade The expanded scope for electronic delivery of the purchase emphasized by Bacchetta et al (1998) raises several issues of particular concern for the WTO. First, there is the issue of whether the transaction involves a good or a service. This issue arises, for example, when considering a good that is delivered from a foreign source via digital instructions for additive manufacturing ("3-D printing") in a domestic location. This issue is consequential because of the different structure and level of market access commitments in GATT (which applies to traded goods) and GATS (which applies to traded services), a difference that reflects in part the different policy instruments available to governments for influencing trade flows across the goods and services sectors (see Staiger, 2018, and Staiger and Sykes, 2021).

Second, if the transaction is determined to involve a service, there is the issue of by what mode the service is being delivered. GATS distinguishes among four modes of trade. "Mode 1" trade involves the cross-border sale of a service from the exporting country to a consumer in the importing country, while "mode 2" involves the consumption of a service in the exporting country by a national of another importing country. "Mode 3" trade involves the establishment of a commercial presence in the importing nation by a foreign service provider. And "mode 4" trade occurs when a foreign supplier not only establishes a commercial presence in the importing nation, but also employs foreign nationals in its operations. Because WTO member governments have typically scheduled different levels of market access commitments across the various modes of trade, the assignment of services to these modes is consequential.

The issues raised by Bacchetta et al (1998) have been debated ever since. Here is how Azmeh, Foster and Echavarri (2020) characterize one way in which this debate is playing out today:

> Historically, the international trade regime has paid more attention to mode 1, reflecting the fact that it is easier to regulate and the most common form of delivery. As such, states impose rules on products exported to their markets, including tariffs and standards, but have fewer restrictions on their citizens traveling abroad to consume services and buy goods (including bringing these goods home in a small scale). In the GATS agreement, states tend to provide less restrictive market access on consumption abroad in comparison to cross-border delivery.
>
> Digital trade predominantly defined under mode 1 would provide states with higher ability to determine which digital goods and services a consumer might access. Thus, they might have the ability to control the conditions, standards, and regulations that providers of those digital goods and services have to follow. The challenge facing policy makers here is operational—how to expand the application of existing rules into these new types of trade flows, including small scale trade through digital platforms.
>
> In contrast, thinking of digital trade as mode 2 entails that through the internet the ability to travel abroad and consume have expanded from a relatively small percentage of the population (with access to passports and financial means) into every citizen with internet access. Consumers visiting foreign websites to purchase goods and services is akin to them traveling abroad physically. In such an understanding of digital trade, the role of the state in shaping market access and the standards and regulations, at least using traditional tools, becomes limited. ... (p 676)

A taxonomy For the purposes of my formal analysis below, I need to partition the kinds of digital trades described above into "digital trade in goods" and "digital trade in services." To this end I will say that trade is "digital" if it involves digital elements in any of the three stages of search, order and payment, or delivery, as described by Bacchetta et al (1998). And I will define a transaction as involving a "good" ("service") if at the moment of consumption that transaction is a good (service) as traditionally defined, i.e., as defined in the pre-digital world.[9]

[9] That said, it should be noted that the exact boundary between goods and services is nowhere defined in the WTO (see Smith and Woods, 2005).

Thus, for my purposes, an Internet search, order and payment that results in the physical delivery of a book from a foreign publisher would be classified as digital trade in a good. And the cross-border transmission of instructions for 3-D printing of a wallet will be classified as digital trade of a good (because it is a wallet at the moment of consumption), while the cross-border purchase of an e-book will be classified as digital trade of a service (because at the moment of consumption the purchaser is not acquiring the book, but rather the licence to read it). And some transactions, such as the importation of a car with Internet connectivity, may involve digital trade in both a good and a service.

Where it is relevant below, I will adopt an analogous approach to identifying modes of supply within services: if a service was classified as mode X in the pre-digital world, then the digital service continues to be classified as mode X. Thus, if in the pre-digital world a domestic citizen would have traveled abroad and taken an in-person foreign cooking class but in the digital world stays at home and enrolls in the cooking class provided virtually from the foreign country over the Internet, I will continue to classify this in the digital world as trade in a mode-2 service, just as it would have been classified in the pre-digital world.

What about truly "new" goods or services that did not exist in the pre-digital world? For the purposes of my taxonomy these goods and services could be treated like any new good in GATT or new service in GATS, that is, initially placed in the "other" category of the internationally standardized Harmonized System (HS) used by the WTO to classify goods and services.[10]

Importantly, in the formal analysis to follow I will assume that digitalization does not itself alter the feasibility of imposing a tariff on the trade in question: if a tariff can be imposed on the physical importation of a wallet, then I am assuming it is also feasible to impose the tariff on the imported wallet when it is delivered by the cross-border transmission of instructions for 3-D printing. In section 6 I consider the possibility that this assumption does not hold, and in light of my analytical findings I propose there a possible modification to the classification of digital trade in goods and services that takes a functional approach to the classification question.

[10] Though see Bacchetta et al (1998, p 51) for a discussion of possible issues with this approach when it comes to new services, and see WTO (2018, pp 168-170) for a review of WTO disputes that centered on the legal question of what qualifies as a "new" good or service.

3. What are the Policies that Impact Digital Trade?

I noted in the Introduction that I will take a reduced-form approach to modeling the impacts of digitalization on international trade. In particular and as noted there, in the analytical sections that follow I will simply assume that digitalization reduces the costs of international trade, and that in choosing their digital trade policies governments weigh the trade-cost reducing impacts that a more open digital environment may bring against any possible non-pecuniary externalities that may arise with such a policy. But what are these digital policies? In this section I describe briefly the kinds of underlying policies that I attempt to capture with my reduced form modeling approach.

Tariffs Tariffs are the most straightforward policy available to governments for restricting trade, and in principle tariffs could constitute a central policy instrument for at least some kinds of digital trade, just as they have for trade in goods historically. But even here the impact of digitalization would be felt, as the small order sizes that the Internet has enabled might necessitate changes in the de minimis values below which customs duties do not apply (see, for example, USITC, 2013, p 5-23). More importantly, WTO members agreed in 1998 to a moratorium on customs duties applied to "electronic transmissions of digital products and services," and this moratorium has been renewed every two years ever since.[11] That said, the issue of exactly what kinds of digital trade this moratorium applies to, and whether the moratorium should be renewed, strengthened to a permanent commitment, or abandoned has been a growing source of contention among WTO members as the importance of digital trade has grown (see Azmeh, Foster and Echavarri, 2020). The future role of tariffs as a policy for restricting digital trade is therefore uncertain.

Regulatory barriers to digital trade Be that as it may, the more novel and contentious policies that impact digital trade involve a variety of regulatory issues. As a general matter there are of course many regulations that have a restrictive impact on international trade and yet serve a legitimate public policy purpose, making the distinction between "protection" and "protectionism" often difficult to draw; the regulatory issues that arise with digital trade are no exception in this regard. But in the digital sphere many of the regulatory issues are still quite novel

[11] The 2019 renewal of the WTO moratorium on customs duties applied to electronic transmissions can be found at https://www.wto.org/english/news_e/news19_e/gc_10dec19_e.htm.

and lie at the center of concerns over highly sensitive questions regarding privacy, law enforcement and national security, where shared norms of best practice and reasonable behavior are not yet fully developed or broadly accepted. Indeed, as Aaronson (2018) notes:

> Many allegations of digital protectionism are concerns about different approaches to regulating the data flows that underpin the internet within national borders. Although the United States and the EU are trying to create shared rules, the two trade giants have also been the most vociferous in describing other countries' approaches as 'protectionist'. (p 3)

Nevertheless, there are some clear regulatory policy areas with respect to digital trade where disciplines in the context of trade agreements are likely to be relevant. Azmeh, Foster and Echavarri (2020, pp 677-678) provide a useful discussion in this regard (see also Meltzer, 2013). They offer as illustration three examples of such policies: internet filtering to block cross-border access to certain websites, data localization, and source-code transfer requirements. As they note:

> Access to websites, digital tools, and services located on foreign servers is a prerequisite to access digital goods and services provided by firms, including both digitally delivered products and physically delivered products. As such, engineering the structure of the internet to block such cross-border access is a very effective way of controlling digital trade.
>
> ... Data localization policy is used to control trade flows and access to foreign digital products. Data localization includes a number of policies that demand that data (or certain categories of data) generated within a state are subject to additional rules, typically rules requiring the storage of data domestically. Such a policy raises the cost of global firms serving a market by demanding that foreign digital firms build or purchase domestic data storage capacities. Through such a policy, data localization could strengthen the position of domestic firms and strengthen local digital ecosystems.
>
> ... Often as part of security requirements, a number of states adopt policies that seek to mandate technology transfer through policies such as source code transfer requirements. Such conditions can have major economic implications, as most companies will consider access to their source code a red line (due to the risks of losing key intellectual property), leading to this requirement serving as a market access restriction. Parallel policies

> are also emerging in regard to mandating firms to reveal encryption keys and algorithms. This can be a major issue for companies, as it could lead to blocking market access if they refuse to comply or to jeopardizing data security and the trust of customers if they do. (Azmeh, Foster and Echavarri, 2020, pp 677-678).

The reduced form approach to modeling the impacts of digitalization on international trade that I introduce in the sections below cannot, of course, do justice to these subtle and complex policy issues, but neither is that my intent. Instead, motivated by these kinds of digital trade policies and the issues that they raise, my intent is simply to represent in a simple and tractable model some of the basic tradeoffs that governments must confront when making their digital policy choices, namely, tradeoffs between the lower trade costs that will obtain with a more open digital policy environment and the various non-trade issues that will arise.

With this in mind, I now turn to developing a pair of benchmark models that can help to illuminate the purpose of trade agreements in a pre-digital and digital world. I begin in the next section by considering the case of digital trade in goods.

4. Digital Trade in Goods[12]

In this section I focus on the case of digital trade in goods. I will think of an open digital policy environment as contributing to lower trade costs but also possibly generating an externality (e.g., related to privacy issues or concerns about national security) that may have both local and cross-border dimensions. I first develop the benchmark model for trade in goods at a general level that incorporates both the pre-digital world and the digital world as special cases. I then consider the purpose of a trade agreement in each special case of this benchmark model.

4.1. A Benchmark trade-in-goods model

I consider a simple partial equilibrium setting, in which a home country imports a competitively produced good from the foreign country, and I let $I \in [0, \infty)$ and $I^* \in [0, \infty)$ denote respectively the home and foreign digital ("Internet") policies, with $I = 0$ $(I^* = 0)$ corresponding to the absence of a workable Internet in the

[12] The material in this section is a slightly adapted and reinterpreted version of the material in section 4.1.1 of Staiger (2019).

home (foreign) country, and with a higher level of I (I^*) corresponding to a more open digital policy environment in the home (foreign) country. I assume that these policies, which I take to represent the kinds of regulatory barriers to digital trade described in section 3, jointly determine the efficiency of trade transactions between the two countries, and in particular that the per-unit (specific) trade cost for exports from foreign to home, ι, can be represented by the function $\iota(I, I^*)$, where $\iota(0,0)$ is non-prohibitive and with $\iota(I, I^*)$ decreasing and convex in both its arguments and non-negative for all I and I^*.[13] To fix ideas, the good under consideration could be a wallet, and the openness of digital policies in each country could determine the efficacy of digital search including the possibility of targeted adds, the functioning of digital payment systems, and the feasibility of digital delivery through 3D printing.

With the import tariff set by the home government (and expressed in specific terms) denoted by τ, and the export tax set by the foreign government (and expressed in specific terms) denoted by τ^*, the arbitrage relationship between the home-country price of this good (P) and the foreign-country price of the good (P^*) that must hold as long as strictly positive trade occurs is given by

$$P = P^* + \iota(I, I^*) + \tau + \tau^*. \tag{4.1}$$

I define the *foreign world price* and the *home world price* by $P^{w*} \equiv P^* + \tau^*$ and $P^w \equiv P - \tau$, respectively. The foreign and home world prices are measures of the foreign- and home-country terms of trade – the foreign terms of trade improves when P^{w*} rises, and the home terms of trade improves when P^w falls – and through (4.1) they are related by $P^w - P^{w*} = \iota(I, I^*)$. A drop in trade costs ι brings P^w and P^{w*} closer together, and when $\iota = 0$ the home and foreign world prices are equated.

I denote the home and foreign demands for the product under consideration by $D(P)$ and $D^*(P^*)$, and I assume that these demand functions are decreasing in their arguments; and for simplicity I assume that the product is supplied only by the foreign country, and denote foreign supply by the increasing function

[13] Rather than reducing the per-unit costs of trade, a more realistic assumption might be that the Internet reduces the fixed cost of entering a particular market, as in Freund and Weinhold (2004). Incorporating such an assumption would require the use of a more involved model of (firm-level) trade than the simple (industry-level) model I adopt here, but I suspect that my results would extend to that setting (see for example Bagwell and Lee, 2020, who show that the purpose of a trade agreement is unchanged by the introduction of heterogeneous firms).

$S^*(P^*)$.[14] Using the pricing relationship (4.1), and denoting foreign export supply by $E^*(P^*) \equiv S^*(P^*) - D^*(P^*)$ and home import demand by $M(P) \equiv D(P)$, the market clearing condition may be written as

$$M(P^* + \iota(I, I^*) + \tau + \tau^*) = E^*(P^*),$$

yielding the market clearing foreign price $\hat{P}^*(\iota(I, I^*) + \tau + \tau^*)$ from which the market clearing home price and foreign and home world prices also follow:

$$\begin{aligned}
\hat{P}(\iota(I, I^*) + \tau + \tau^*) &\equiv \hat{P}^*(\iota(I, I^*) + \tau + \tau^*) + \iota(I, I^*) + \tau + \tau^* \\
\hat{P}^{w*}(\iota(I, I^*) + \tau, \tau^*) &\equiv \hat{P}^*(\iota(I, I^*) + \tau + \tau^*) + \tau^* \\
\hat{P}^{w}(\iota(I, I^*) + \tau^*, \tau) &\equiv \hat{P}(\iota(I, I^*) + \tau + \tau^*) - \tau.
\end{aligned}$$

As is standard, the world prices depend on the levels of both τ and τ^*, but the home and foreign prices depend only on the sum $\tau + \tau^*$ (and on the digital policies I and I^*).

With the market clearing price expressions above, the terms-of-trade impacts of policy choices can now be assessed. Regarding the terms-of-trade impacts of trade taxes, direct calculations yield (with a prime denoting the derivative of the function with respect to its argument):

$$\begin{aligned}
\frac{\partial \hat{P}^w}{\partial \tau} &= \frac{\partial \hat{P}^{w*}}{\partial \tau} = \frac{M'}{E^{*\prime} - M'} < 0 \\
\frac{\partial \hat{P}^{w*}}{\partial \tau^*} &= \frac{\partial \hat{P}^{w}}{\partial \tau^*} = \frac{E^{*\prime}}{E^{*\prime} - M'} > 0.
\end{aligned}$$

As expected, an increase in the home-country tariff improves the home terms of trade and worsens the foreign terms of trade, while an increase in the foreign-country tariff has the opposite impact, improving the foreign terms of trade and worsening the home terms of trade. These familiar terms-of-trade effects of tariff intervention provide the basis for the inefficient Prisoners' Dilemma situation that according to the terms-of-trade theory of trade agreements arises in the absence of a trade agreement.[15]

[14] Here and throughout I will I assume that supply curves are unaffected by digital policy, but none of the conclusions I emphasize below depend on this assumption.

[15] See Johnson (1953-54) and Bagwell and Staiger (1999, 2002). I follow Bagwell and Staiger's terminology in referring to theories of trade agreements in which the terms-of-trade effects of policies play a central role in determining the purpose of trade agreements as "terms-of-trade" theories. Staiger (forthcoming) provides a recent assessment of this theory.

The terms-of-trade impacts of digital openness (increases in I and I^*) are more novel.[16] The terms-of-trade impacts of a more open home-country digital policy are given by

$$\begin{aligned} \frac{\partial \hat{P}^w}{\partial \iota}\frac{\partial \iota}{\partial I} &= \frac{E^{*\prime}}{E^{*\prime} - M^{\prime}} \times \frac{\partial \iota}{\partial I} < 0 \\ \frac{\partial \hat{P}^{w*}}{\partial \iota}\frac{\partial \iota}{\partial I} &= \frac{M^{\prime}}{E^{*\prime} - M^{\prime}} \times \frac{\partial \iota}{\partial I} > 0, \end{aligned} \tag{4.2}$$

while for a more open foreign-country digital policy these impacts are given by

$$\begin{aligned} \frac{\partial \hat{P}^{w*}}{\partial \iota}\frac{\partial \iota}{\partial I^*} &= \frac{M^{\prime}}{E^{*\prime} - M^{\prime}} \times \frac{\partial \iota}{\partial I^*} > 0 \\ \frac{\partial \hat{P}^w}{\partial \iota}\frac{\partial \iota}{\partial I^*} &= \frac{E^{*\prime}}{E^{*\prime} - M^{\prime}} \times \frac{\partial \iota}{\partial I^*} < 0. \end{aligned} \tag{4.3}$$

Evidently, an open home-country digital policy improves the home-country terms of trade *while at the same time improving the terms of trade of the foreign country*, and similarly for an open foreign-country digital policy.

This "win-win" prospect for open digital policies makes it tempting to conclude that the terms-of-trade theory cannot explain why countries would need an international agreement to encourage open digital policies at all. But as I will demonstrate below, this is not correct. Intuitively, the key is to note from the derivative expressions in (4.2) and (4.3) above that each country's digital openness imparts a positive terms-of-trade externality on the other country, providing a possible reason for *under*-investment in digital openness from an international perspective when countries are guided only by their unilateral interests (i.e., in the absence of an international agreement that covers digital trade policies), and therefore a possible role for a trade agreement that encourages digital openness.

I now define the welfare functions for the home and foreign countries. I abstract from political economy/distributional motives, though the results I report below are easily generalized to include such motives.[17] But I allow an open digital environment in a country to generate a non-pecuniary externality, possibly relating to privacy issues (of the kind highlighted, for example, in Acemoglu et

[16] Staiger (2019) makes the same point about investments in trade facilitation more generally.

[17] See Bagwell and Staiger (1999, 2002) on the wide variety of government objectives that can be incorporated into the terms-of-trade theory of trade agreements without altering the theory's prediction about the purpose of a trade agreement.

al, forthcoming) or to national security concerns (as, for example, have been expressed by the United States in the context of the Chinese telecommunications maker Huawei). I assume that this externality takes an "eyesore" form that does not itself affect production and is not internalized by individual consumers and hence does not affect demands, but that detracts in a separable way from aggregate welfare. Moreover, I assume that this externality may have both a local and a cross-border component, with the local component increasing in own-digital openness and the cross-border component increasing in the digital openness of one's trading partner. For example, more digital openness in the home country (higher I) might create a digital environment where malicious software could more easily be downloaded to the devices of home-country citizens; and this malicious software might (or might not) then negatively impact foreign citizens as well.

Formally, I denote by $c(I)$ and $c^*(I^*)$ the local externality generated in the home and foreign country by the respective levels of digital openness I and I^*, with $c(0) = 0$ and $c^*(0) = 0$ and with c and c^* increasing and convex in their respective arguments; and I let the parameter $\theta \in [0, 1]$ govern the degree of cross-border spillovers, so that the externality from the digital policies I and I^* reduces home-country welfare by the amount $[c(I) + \theta c^*(I^*)]$ and reduces foreign-country welfare by the amount $[c^*(I^*) + \theta c(I)]$.

With no home-country production, home welfare is then given by the sum of consumer surplus plus tariff revenue minus the cost of the externality from digital openness. Letting CS denote home-country consumer surplus and using $\tau = P - P^w$, home welfare is given by

$$\begin{aligned} W &= CS(\hat{P}(\iota(I, I^*) + \tau + \tau^*)) \\ &\quad +[\hat{P}(\iota(I, I^*) + \tau + \tau^*) - \hat{P}^w(\iota(I, I^*) + \tau^*, \tau)] \times M(\hat{P}(\iota(I, I^*) + \tau + \tau^*)) \\ &\quad -[c(I) + \theta c^*(I^*)] \\ &\equiv W(I, I^*, \hat{P}(\iota(I, I^*) + \tau + \tau^*), \hat{P}^w(\iota(I, I^*) + \tau^*, \tau)). \end{aligned}$$

Taking account of production in the foreign country and with PS^* denoting foreign producer surplus, foreign welfare is similarly defined as the sum of consumer and producer surplus plus export tax revenue minus the cost of the externality from digital openness, or

$$\begin{aligned} W^* &= CS^*(\hat{P}^*(\iota(I, I^*) + \tau + \tau^*)) + PS^*(\hat{P}^*(\iota(I, I^*) + \tau + \tau^*)) \\ &\quad +[\hat{P}^{w*}(\iota(I, I^*) + \tau, \tau^*) - \hat{P}^*(\iota(I, I^*) + \tau + \tau^*)] \times E^*(\hat{P}^*(\iota(I, I^*) + \tau + \tau^*)) \\ &\quad -[c^*(I^*) + \theta c(I)] \\ &\equiv W^*(I^*, I, \hat{P}^*(\iota(I, I^*) + \tau + \tau^*), \hat{P}^{w*}(\iota(I, I^*) + \tau, \tau^*)). \end{aligned}$$

Finally, the sum of home and foreign welfare, which I refer to as "world welfare" and denote by W^w, is given by

$$\begin{aligned} W^w &= CS(\hat{P}(\iota(I,I^*)+\tau+\tau^*)) + CS^*(\hat{P}^*(\iota(I,I^*)+\tau+\tau^*)) + PS^*(\hat{P}^*(\iota(I,I^*)+\tau+\tau^*)) \\ &\quad +[\hat{P}(\iota(I,I^*)+\tau+\tau^*) - \hat{P}^*(\iota(I,I^*)+\tau+\tau^*) - \iota(I,I^*)] \times E^*(\hat{P}^*(\iota(I,I^*)+\tau+\tau^*)) \\ &\quad -[c(I)+\theta c^*(I^*)] - [c^*(I^*)+\theta c(I)] \\ &\equiv W^w(I,I^*,\hat{P}(\iota(I,I^*)+\tau+\tau^*),\hat{P}^*(\iota(I,I^*)+\tau+\tau^*)). \end{aligned}$$

Notice that while home and foreign welfare each depend on their respective world prices and hence on the levels of both τ and τ^*, world welfare is independent of world prices – because movements in these prices only serve to redistribute surplus between the home and foreign country – and hence world welfare depends only on the sum of home and foreign tariffs $\tau+\tau^*$ (in addition to the digital openness levels I and I^*).

4.2. The purpose of GATT in a pre-digital world

I will think of the pre-digital world as corresponding to a special case of the benchmark trade-in-goods model described above in which

$$I \equiv 0 \equiv I^* \tag{4.4}$$

and there is no workable Internet in either country. In this pre-digital world, the benchmark model is very simple and the purpose of a trade agreement is transparent and familiar. Nevertheless, it is useful to review the steps in determining the purpose of a trade agreement in this setting, so that a comparison with the digital world can be drawn.

Under (4.4), trade costs are exogenous and given by $\iota(0,0) \equiv \bar{\iota}$, and with no workable Internet anywhere in the world we also have an absence of digital externalities $c(0) = c^*(0) = 0$. Hence, in the pre-digital world, the only policy choice to be made by each country is its tariff choice, and the country and world welfare levels as functions of these policy choices are given by

$$\begin{aligned} W &= CS(\hat{P}(\bar{\iota}+\tau+\tau^*)) \\ &\quad +[\hat{P}(\bar{\iota}+\tau+\tau^*) - \hat{P}^w(\bar{\iota}+\tau^*,\tau)] \times M(\hat{P}(\bar{\iota}+\tau+\tau^*)) \\ &\equiv W(\hat{P}(\bar{\iota}+\tau+\tau^*),\hat{P}^w(\bar{\iota}+\tau^*,\tau)), \end{aligned}$$

$$\begin{aligned} W^* &= CS^*(\hat{P}^*(\bar{\iota}+\tau+\tau^*)) + PS^*(\hat{P}^*(\bar{\iota}+\tau+\tau^*)) \\ &\quad +[\hat{P}^{w*}(\bar{\iota}+\tau,\tau^*) - \hat{P}^*(\bar{\iota}+\tau+\tau^*)] \times E^*(\hat{P}^*(\bar{\iota}+\tau+\tau^*)) \\ &\equiv W^*(\hat{P}^*(\bar{\iota}+\tau+\tau^*),\hat{P}^{w*}(\bar{\iota}+\tau,\tau^*)), \text{ and} \end{aligned}$$

$$\begin{aligned} W^w &= CS(\hat{P}(\bar{\iota}+\tau+\tau^*)) + CS^*(\hat{P}^*(\bar{\iota}+\tau+\tau^*)) + PS^*(\hat{P}^*(\bar{\iota}+\tau+\tau^*)) \\ &\quad +[\hat{P}(\bar{\iota}+\tau+\tau^*) - \hat{P}^*(\bar{\iota}+\tau+\tau^*) - \bar{\iota}] \times E^*(\hat{P}^*(\bar{\iota}+\tau+\tau^*)) \\ &\equiv W^w(\bar{\iota}, \hat{P}(\bar{\iota}+\tau+\tau^*), \hat{P}^*(\bar{\iota}+\tau+\tau^*)). \end{aligned}$$

Efficient Policies I define efficient policies as those that maximize world welfare (and thereby implicitly assume that lump sum transfers are available to distribute surplus across the two countries as desired). As noted above, world welfare depends on the sum of the home and foreign tariffs, $\tau + \tau^*$. The first-order conditions that define the sum of efficient tariffs, $\partial W^w/\partial[\tau + \tau^*] = 0$, can be simplified to yield[18]

$$[\tau + \tau^*] \times \frac{\partial E^*}{\partial P^*} \frac{\partial \hat{P}^*}{\partial[\tau + \tau^*]} = 0$$

which immediately implies

$$\tau^e + \tau^{*e} = 0 \tag{4.5}$$

where a superscript "e" denotes efficient policies. Hence, as should come as no surprise in this perfectly competitive setting, there is no efficiency role for tariffs.

Nash Policies Next consider the Nash policies adopted by the two countries in the absence of a trade agreement. The first-order condition for the home country that defines its best-response level of τ is given by

$$\frac{\partial W}{\partial \tau} = -M(\hat{P})\frac{\partial \hat{P}}{\partial \tau} + \tau \frac{\partial E^*}{\partial P^*} \frac{\partial \hat{P}^*}{\partial \tau} + M(\hat{P}) = 0. \tag{4.6}$$

Similarly, the first-order condition for the foreign country that defines its best-response level of τ^* is given by

$$\frac{\partial W^*}{\partial \tau^*} = -E^*(\hat{P}^*)\frac{\partial \hat{P}^*}{\partial \tau^*} + \tau^* \frac{\partial M}{\partial P} \frac{\partial \hat{P}}{\partial \tau^*} + E^*(\hat{P}^*) = 0. \tag{4.7}$$

The Nash policies, which I denote by τ^N and τ^{*N}, satisfy the two first-order conditions in (4.6) and (4.7) simultaneously. It follows that the Nash tariffs are characterized by

$$\tau^N = \frac{\hat{P}^{w*N}}{\eta^{E^{*N}}} \quad \text{and} \quad \tau^{*N} = \frac{\hat{P}^{wN}}{\eta^{M^N}}, \tag{4.8}$$

[18] I assume here and throughout that second-order conditions hold.

with $\eta^{E^{*N}}$ the elasticity of foreign export supply evaluated at Nash policies and η^{M^N} the elasticity of home import demand (defined positively) evaluated at Nash policies, and where $\hat{P}^{w*N}$ and $\hat{P}^{wN}$ are evaluated at Nash policies. The Nash tariffs in (4.8) represent the usual inverse-trade-elasticity formulae for the Johnson (1953-54) optimal tariff.

The purpose of a trade agreement Hence, as has been know since Johnson (1953-54) and as a comparison of (4.8) with (4.5) confirms, when countries have market power in the international markets on which they trade and therefore face finite trade elasticities from abroad, Nash tariffs are too high and Nash trade volumes are too low; and the purpose of a trade agreement in settings such as this is to eliminate the beggar-thy-neighbor incentives that each country has to manipulate the terms-of-trade with its unilateral tariff choice, and thereby to enjoy the mutual benefits from the expanded trade volumes that come from implementation of internationally efficient policies.[19]

Moreover, while for simplicity I have adopted a model where there are no non-tariff policy choices to be made in the pre-digital world, Bagwell and Staiger (2001) establish that the purpose of a trade agreement is unchanged when domestic policies are added to the picture. Specifically, allowing governments to choose domestic (e.g., labor or environmental) standards as well as tariffs, and defining "market access" to reflect the position of a country's import demand curve, Bagwell and Staiger show that in the absence of international negotiations governments would set their tariffs inefficiently high to reduce market access to inefficiently low levels but would choose their domestic standards efficiently *given* these levels of market access: the only problem would be a market-access problem, the same problem whose solution lies at the core of the purpose of a trade agreement in the simple model above.[20]

The design of GATT As Bagwell and Staiger (2001) observe, these findings can illuminate the logic of GATT's design. According to these findings, a trade

[19] Broda, Limao and Weinstein (2008) provided the first systematic evidence that countries possess market power on world markets and use it in setting their non-cooperative tariffs. See Bagwell, Bown and Staiger (2016) for a recent review of the terms-of-trade theory of trade agreements and the empirical literature on its main tenets.

[20] See Bagwell and Staiger (2002) for further development of this definition of market access and its link to the terms-of-trade theory, and to the "conditions of competition for trade" that defines market access in the WTO.

agreement might focus on lowering tariffs as a means of expanding market access and trade volumes to efficient levels, while putting in place various "market access preservation rules" that apply to non-tariff (e.g., behind-the-border) policies and prevent governments from back-sliding on the market access commitments implied by their negotiated tariff bindings with new protective non-tariff measures. In principle, a trade agreement designed in this way would be capable of bringing countries to the efficiency frontier. At a broad level, this logic fits nicely with the basic structure of GATT's shallow integration approach.[21]

Bagwell and Staiger (1999, 2001, 2002, 2016) show that these findings extend to a variety of economic settings and diverse government policy preferences. I consider next whether they hold in a world of digital trade in goods.

4.3. The purpose of GATT in a digital world

In a world of digital trade in goods, the benchmark model of section 4.1 applies. As I did when considering the pre-digital world, below I first characterize efficient policies, then characterize Nash policies, and then consider the purpose of a trade agreement covering goods trade in a digital world. It is useful to consider two possibilities for the digital world, one where the non-pecuniary externality associated with digital openness is purely local and does not spill over across borders, and the other where this externality has both local and cross-border components.

4.3.1. A purely local non-pecuniary externality from digital openness

I consider first the case of a purely local non-pecuniary externality from digital openness. To this end, I now impose for the remainder of this section the following

$$\textit{No cross-border non-pecuniary externality: } \theta \equiv 0, \qquad \text{(Assumption 1)}$$

and I record the country and world welfare levels in a digital world under Assumption 1:

$$\begin{aligned} W &= CS(\hat{P}(\iota(I,I^*)+\tau+\tau^*)) \\ &\quad +[\hat{P}(\iota(I,I^*)+\tau+\tau^*)-\hat{P}^w(\iota(I,I^*)+\tau^*,\tau)]\times M(\hat{P}(\iota(I,I^*)+\tau+\tau^*))-c(I) \\ &\equiv W(I,\hat{P}(\iota(I,I^*)+\tau+\tau^*),\hat{P}^w(\iota(I,I^*)+\tau^*,\tau)). \end{aligned}$$

[21] According to this approach, countries use their tariff negotiations to make market access commitments and the central role of GATT Articles is then to provide the accompanying "market access preservation rules." This role of GATT Articles is reflected, for example, in Petersmann's (1997, p. 136) observation that "...the function of most GATT rules (such as Articles I-III and XI) is to establish conditions of competition and to protect trading opportunities...".

$$
\begin{aligned}
W^* &= CS^*(\hat{P}^*(\iota(I,I^*)+\tau+\tau^*)) + PS^*(\hat{P}^*(\iota(I,I^*)+\tau+\tau^*)) \\
&\quad +[\hat{P}^{w*}(\iota(I,I^*)+\tau,\tau^*) - \hat{P}^*(\iota(I,I^*)+\tau+\tau^*)] \times E^*(\hat{P}^*(\iota(I,I^*)+\tau+\tau^*)) - c^*(I^*) \\
&\equiv W^*(I^*, \hat{P}^*(\iota(I,I^*)+\tau+\tau^*), \hat{P}^{w*}(\iota(I,I^*)+\tau,\tau^*)).
\end{aligned}
$$

$$
\begin{aligned}
W^w &= CS(\hat{P}(\iota(I,I^*)+\tau+\tau^*)) + CS^*(\hat{P}^*(\iota(I,I^*)+\tau+\tau^*)) + PS^*(\hat{P}^*(\iota(I,I^*)+\tau+\tau^*)) \\
&\quad +[\hat{P}(\iota(I,I^*)+\tau+\tau^*) - \hat{P}^*(\iota(I,I^*)+\tau+\tau^*) - \iota(I,I^*)] \times E^*(\hat{P}^*(\iota(I,I^*)+\tau+\tau^*)) \\
&\quad -c(I) - c^*(I^*) \\
&\equiv W^w(I, I^*, \hat{P}(\iota(I,I^*)+\tau+\tau^*), \hat{P}^*(\iota(I,I^*)+\tau+\tau^*)).
\end{aligned}
$$

Efficient Policies As before, I define efficient policies as those that maximize world welfare. In the digital world, world welfare depends on the sum of the home and foreign tariffs, $\tau+\tau^*$, and on the degree of home and foreign digital openness, I and I^*. As before, the first-order conditions that define the sum of efficient tariffs, $\partial W^w/\partial[\tau+\tau^*]=0$, can be simplified to yield

$$[\tau+\tau^*] \times \frac{\partial E^*}{\partial P^*}\frac{\partial \hat{P}^*}{\partial[\tau+\tau^*]} = 0$$

which immediately implies

$$\tau^e + \tau^{*e} = 0. \tag{4.9}$$

Hence, again there is no efficiency role for tariff intervention, and this is true independent of the degree of digital openness in each country (and hence independent of trade costs ι).

Consider next the efficient level of home and foreign digital openness, denoted by I^e and I^{*e} respectively. The first-order condition that defines I^e can be manipulated to yield

$$\{[\tau+\tau^*] \times \frac{\partial E^*}{\partial P^*}\frac{\partial \hat{P}^*}{\partial[\tau+\tau^*]} - E^*\}\frac{\partial \iota}{\partial I} = c'(I^e)$$

which, evaluated at the efficient tariffs $\tau^e+\tau^{*e}$, simplifies to

$$M^e \times [-\frac{\partial \iota}{\partial I}] = c'(I^e) \tag{4.10}$$

where M^e denotes home import volume evaluated at efficient policies and where in writing (4.10) I have used the market clearing condition $M=E^*$. In words, the efficient level of home-country digital openness I^e equates the marginal benefit of the last unit of digital opening allowed by the home country (the marginal savings in total trade costs $M^e \cdot [-\frac{\partial \iota}{\partial I}]$) with the marginal cost to the home country owing to the impact of this additional digital opening on the local non-pecuniary externality that is associated with the home country's digital openness ($c'(\cdot)$). The efficient level of digital openness for the foreign country, I^{*e}, is similarly characterized:

$$M^e \times [-\frac{\partial \iota}{\partial I^*}] = c^{*\prime}(I^{*e}). \tag{4.11}$$

Notice from (4.10) and (4.11) that, if home-country and foreign-country digital openness enter symmetrically into the trade cost function $\iota(I, I^*)$, then the efficient digital policies will differ across countries to the extent that the (local) externalities from digital openness differ in the two countries. Hence, the functions $c(\cdot)$ and $c^*(\cdot)$ play a role analogous to the difference across countries in their "preferences" for labor or environmental standards entertained by Bagwell and Staiger (2001).[22]

Nash Policies Next consider the Nash policies adopted by the two countries in the absence of a trade agreement. The first-order conditions for the home country that define its best-response levels of τ and I are given by

$$\begin{aligned} \frac{\partial W}{\partial \tau} &= -M(\hat{P})\frac{\partial \hat{P}}{\partial \tau} + \tau \frac{\partial E^*}{\partial P^*}\frac{\partial \hat{P}^*}{\partial \tau} + M(\hat{P}) = 0 \\ \frac{\partial W}{\partial I} &= [-M(\hat{P})\frac{\partial \hat{P}}{\partial \iota} + \tau \frac{\partial E^*}{\partial P^*}\frac{\partial \hat{P}^*}{\partial \iota}]\frac{\partial \iota}{\partial I} - c'(\cdot) = 0. \end{aligned} \tag{4.12}$$

Similarly, the first-order conditions for the foreign country that define its best-response levels of τ^* and I^* are given by

$$\begin{aligned} \frac{\partial W^*}{\partial \tau^*} &= -E^*(\hat{P}^*)\frac{\partial \hat{P}^*}{\partial \tau^*} + \tau^* \frac{\partial M}{\partial P}\frac{\partial \hat{P}}{\partial \tau^*} + E^*(\hat{P}^*) = 0 \\ \frac{\partial W^*}{\partial I^*} &= [-E^*(\hat{P}^*)\frac{\partial \hat{P}^*}{\partial \iota^*} + \tau^* \frac{\partial M}{\partial P}\frac{\partial \hat{P}}{\partial \iota^*}]\frac{\partial \iota^*}{\partial I^*} - c^{*\prime}(\cdot) = 0. \end{aligned} \tag{4.13}$$

[22] In this regard, my modeling of digital openness therefore resonates with the position of Goldfarb and Trefler (2019), who argue that a country's choice of privacy policy in the context of Artificial Intelligence technologies raises issues analogous to those associated with the national choice of labor and environmental standards. However, in contrast to Goldfarb and Trefler and to Bagwell and Staiger (2001), I focus here on the impact of digital openness on trade costs.

The Nash policies, which I denote by τ^N, I^N, τ^{*N} and I^{*N}, satisfy the four first-order conditions in (4.12) and (4.13) simultaneously.

Now notice from the pricing relationships above that $\frac{\partial \hat{P}}{\partial \tau} = \frac{\partial \hat{P}}{\partial \iota}$ and $\frac{\partial \hat{P}^*}{\partial \tau} = \frac{\partial \hat{P}^*}{\partial \iota}$ and that $\frac{\partial \hat{P}^*}{\partial \tau^*} = \frac{\partial \hat{P}^*}{\partial \iota^*}$ and $\frac{\partial \hat{P}}{\partial \tau^*} = \frac{\partial \hat{P}}{\partial \iota^*}$. Using this, substituting the top first-order condition in (4.12) into the bottom first-order condition in (4.12), and simplifying the top condition in (4.12) further, and performing the analogous steps for the first order conditions in (4.13), it follows that the Nash tariffs are characterized by

$$\tau^N = \frac{\hat{P}^{w*N}}{\eta^{E^{*N}}} \text{ and } \tau^{*N} = \frac{\hat{P}^{wN}}{\eta^{M^N}}, \tag{4.14}$$

while the Nash digital policies satisfy

$$M^N \times [-\frac{\partial \iota}{\partial I}] = c'(I^N) \text{ and } M^N \times [-\frac{\partial \iota}{\partial I^*}] = c^{*\prime}(I^{*N}), \tag{4.15}$$

where, as before, $\eta^{E^{*N}}$ is the elasticity of foreign export supply evaluated at Nash policies and η^{M^N} is the elasticity of home import demand (defined positively) evaluated at Nash policies, and where $\hat{P}^{w*N}$, $\hat{P}^{wN}$ and M^N denote their respective previously-defined magnitudes evaluated at Nash policies. The Nash tariffs in (4.14) again represent the usual Johnson optimal tariff; and the Nash digital openness levels described by (4.15) equate the marginal benefit of further digital openness with its marginal cost, just as described previously in the context of efficient policy choices.

The purpose of a trade agreement What is the difference between the problem for GATT to solve in the digital world and the insufficient-market-access problem that I characterized in the pre-digital world of the previous section? As in the pre-digital world, Nash tariffs in the digital world are inefficiently high, as a comparison between (4.14) and (4.9) confirms. And in the digital world, the degree of digital openness as reflected in the levels of I and I^* is too low in the Nash equilibrium relative to the efficient level, as can be confirmed from a comparison of (4.15) and (4.10)-(4.11) and noting that $M^N < M^e$ implies $I^N < I^e$ and $I^{*N} < I^{*e}$.[23]

But notice that *given* the Nash level of trade volume M^N, the expressions in (4.15) that implicitly define I^N and I^{*N} are identical to the expressions in (4.10)-(4.11) that implicitly define I^e and I^{*e}. This is the same structure that

[23] This follows from the convexity of $c(\cdot)$ and $c^*(\cdot)$.

Bagwell and Staiger (2001) exploited to demonstrate that the only problem for a trade agreement to solve is the insufficient-market-access problem, the same problem that I characterized in the pre-digital world of the previous section. And it implies that the purpose of a trade agreement is unchanged by the introduction of digital trade in goods. Specifically, in both the pre-digital and digital worlds, the purpose of GATT is to reduce tariffs and thereby expand market access to efficient levels.

The design of GATT in a digital world If design reflects purpose, then the common purpose of GATT across the pre-digital and digital worlds that I have just described suggests that in principle there is no reason for the design of GATT to change in the presence of digital trade in goods. To see what this implies for the design of GATT in a digital world, I now describe how the logic of shallow integration – the same logic that Bagwell and Staiger (2001) showed would hold when governments can chose both tariffs and domestic standards – continues to apply in the digital world I have modeled here.

A first observation is that an agreement to set tariffs at free trade will not by itself achieve the efficiency frontier, because as is easily shown such an agreement would induce each country to reduce its digital openness as a second-best means of manipulating its terms of trade once the agreement constrains its tariffs away from such behavior. One approach, then, is for governments to negotiate over both tariffs and their digital policies to ensure that the conditions for efficiency in (4.9) and (4.10)-(4.11) are met. But as noted above, this deep-integration approach is not the approach that GATT has traditionally taken. Instead, GATT's shallow-integration approach can in principle work in a digital world as follows.

Beginning from the Nash policies characterized by (4.14)-(4.15), suppose that governments negotiate over tariffs and agree to a pair of tariffs $\tilde{\tau}$ and $\tilde{\tau}^*$ that satisfy the condition

$$M(\hat{P}(\iota(I^N, I^{*N}) + \tilde{\tau} + \tilde{\tau}^*)) = M^e. \tag{4.16}$$

According to (4.16), in this first step countries would agree to a pair of tariffs that deliver the efficient level of market access and hence import volume in light of their existing non-cooperative digital policies I^N and I^{*N}. Given that the non-cooperative digital policies are less open than efficient digital policies, (4.16) implies that the tariffs that governments would agree to in their negotiation would be below the efficient free-trade levels.

Then, subsequent to these tariff negotiations, suppose that each government is allowed to set/adjust its digital policy unilaterally, but that governments are subject to a "market access preservation rule" of the following form: if a government alters its digital policy from the Nash level, it must also adjust its tariff to preserve the market access level implied by its original negotiated tariff commitment. Since policy adjustments by one country which preserve market access also preserve the equilibrium trade volume, policy adjustments for the home government that satisfy the market access preservation rule are defined by

$$\frac{d\tau}{dI}|_{dM=0} = \frac{-\frac{\partial \hat{P}}{\partial \iota}\frac{\partial \iota}{\partial I}}{\frac{\partial \hat{P}}{\partial \tau}} > 0 \qquad (4.17)$$

with an analogous condition holding for the foreign government, and with the inequality in (4.17) following from the signs of price derivatives reported above and implying that the home government must lower (can raise) its tariff if it reduces (increases) its digital openness subsequent to tariff negotiations with the foreign country, and similarly for the foreign government.

Subject to this market access preservation rule, what digital policy will the home government choose? Its unilateral choice will satisfy the first-order condition

$$\frac{\partial W}{\partial I} + \frac{\partial W}{\partial \tau}\frac{d\tau}{dI}|_{dM=0} = 0$$

which implies

$$[-M^e \frac{\partial \hat{P}}{\partial \iota} + \tau \frac{\partial E^*}{\partial P^*}\frac{\partial \hat{P}^*}{\partial \iota}]\frac{\partial \iota}{\partial I} - c'(\cdot) + [-M^e \frac{\partial \hat{P}}{\partial \tau} + \tau \frac{\partial E^*}{\partial P^*}\frac{\partial \hat{P}^*}{\partial \tau} + M^e] \times \frac{-\frac{\partial \hat{P}}{\partial \iota}\frac{\partial \iota}{\partial I}}{\frac{\partial \hat{P}}{\partial \tau}} = 0.$$

But using the price derivatives reported above, this first-order condition simplifies to

$$M^e \times [-\frac{\partial \iota}{\partial I}] = c'(I^e),$$

the condition for the efficient home-country digital policy reported in (4.10).

Hence, with the market access preservation rule preventing the home-country government from manipulating the terms of trade with its choice of digital policy, the home-country government would choose to liberalize its digital policy to the efficient level I^e while at the same time raising its tariff from $\tilde{\tau}$ to its efficient level so as to preserve the (efficient) level of market access that was implied by its tariff negotiations. An analogous argument applies for the foreign-country unilateral choice of digital policy. Evidently, by following GATT's shallow-integration approach (see note 21), governments can ensure that the conditions for efficiency in (4.9) and (4.10)-(4.11) are met.

At a conceptual level, the arguments above imply that, despite the fact that digitalization has indisputably permeated deeply into the modern world, there is no more (or less) reason to take a "holistic" (i.e., deep) approach to liberalization in the digital world than there was in the pre-digital world. At a more practical level, these arguments do not so much mean that no degree of deep integration is necessary in a digital world as they provide a potentially useful guardrail that can help delineate the "depth" that deep trade agreements should go in the digital world: according to these results, just as in the pre-digital world there is no reason to go deeper in a trade agreement for the world of digital trade in goods than what is required to ensure that property rights over negotiated market access are reasonably secure. I will return to these considerations in section 6 when I discuss in more detail the implications of digital trade for the design of the WTO.

4.3.2. A cross-border non-pecuniary externality from digital openness

In the previous section I maintained Assumption 1 in order to abstract from the possibility that an open digital policy might lead to cross-border non-pecuniary externalities. I now relax Assumption 1, and allow the non-pecuniary externalities associated with digital trade to cross borders.

Notice that the presence of cross-border non-pecuniary externalities does not alter the Nash policy choices of the two governments, which will still be characterized by (4.14)-(4.15), because with $\theta > 0$ the only difference in a country's welfare function is that its welfare is now reduced directly (for given prices) when its *trading partner's* digital policy is more open, and this is a policy choice that a country has no say in under noncooperative Nash choices. But with $\theta > 0$, the efficient policy choices will now be different than they were under Assumption 1. In particular, it is direct to show that while the efficient tariffs are still characterized by (4.9), the efficient levels of digital openness are now characterized by

$$M^e \times [-\frac{\partial \iota}{\partial I}] = [1+\theta] \times c'(I^e) \tag{4.18}$$

for the home-country government and by

$$M^e \times [-\frac{\partial \iota}{\partial I^*}] = [1+\theta] \times c^{*\prime}(I^{*e}) \tag{4.19}$$

for the foreign-country government. As is intuitive and as a comparison of (4.18)-(4.19) with (4.10)-(4.11) confirms, efficiency requires a less-open digital policy for each country when digital openness imposes cross-border (negative) externalities on trading partners.[24] And this implies in turn that the efficient trade volume will also be lower.

What does this mean for the purpose of a trade agreement? In fact, as I now demonstrate, when digital trade in goods generates cross-border non-pecuniary externalities, the purpose of a trade agreement is no longer simply to solve the insufficient-market-access problem that I characterized in the pre-digital world. Rather, there are now two problems to solve: the cross-border non-pecuniary externality must be addressed; and conditional on addressing this first problem, the insufficient market access problem familiar from the pre-digital world must also be addressed.

To see that the problem is now more complex, suppose first that countries attempted the same shallow-integration approach that I described in the previous section, where Assumption 1 held. By negotiating to the appropriate tariffs, governments could position their market access at the efficient level appropriate for the case where Assumption 1 is violated and $\theta > 0$. But the problem is that under the subsequent unilateral policy adjustments subject to the market access preservation rule, governments would choose levels of digital openness I and I^* that satisfied (4.10)-(4.11), not the conditions for efficiency (4.18)-(4.19) that are relevant when $\theta > 0$. This implies that the GATT's shallow approach to integration is no longer appropriate for a world of digital trade in goods if that trade generates significant cross-border non-pecuniary externalities.

Notice, though, that even in this case there may be an approach to integration suggested by the discussion above that lies somewhere between shallow integration on the one hand, and on the other hand fully deep integration where countries negotiate over all aspects of digital policy that enter into I and I^* in addition to their tariffs. This is because if countries could focus on just those aspects of their digital openness that are generating the cross–border non-pecuniary externalities and address those externalities with limited negotiations, then, from that point forward, we would be back in the world of the previous section where Assumption 1 applies and where achieving the efficiency frontier with GATT's shallow approach to integration is possible.[25] I explore this possibility further in section 6.

[24] Again this follows from the convexity of $c(\cdot)$ and $c^*(\cdot)$.

[25] While the existence of a cross-border non-pecuniary externality that I analyze here therefore does change the purpose of a trade agreement, it does so in a way that suggests the possibility

5. Digital Trade in Services[26]

In this section I focus on the case of digital trade in services. As in the previous section, I will think of an open digital policy environment as contributing to lower trade costs but also possibly generating an externality that may have both local and cross-border dimensions. Mirroring the structure of the previous section, I first develop a benchmark model for services trade at a general level that incorporates both the pre-digital world and the digital world as special cases. I then consider the purpose of a trade agreement in each special case of this benchmark model.

5.1. A benchmark trade-in-services model

I follow Staiger and Sykes (2021) and focus on mode 3 services trade in order to reflect the emphasis of GATS commitments. A good illustrative example might be trade in construction services. In the pre-digital world, a foreign construction company opens a branch office in the home-country market and takes on local road construction projects, thereby providing mode 3 exports of construction services to the home country. In the digital world, all that is still true, but the foreign company advertises its construction services on its foreign website where orders are placed and where electronic payments are also made, and it digitally communicates with the home-country branch office over initial orders and any change orders that arise during the course of construction; and the openness of digital policies in each country can impact the functioning of each of these digital tasks.

As Staiger and Sykes (2021) observe, the approach to market access liberalization taken by GATS is strikingly at odds with the shallow-integration approach taken by GATT, and is instead more aptly described as a deep-integration approach, whereby the negotiated change or removal of domestic regulations and other non-tariff barriers to trade in service sectors is seen as the primary method of expanding market access. And while mode 3 services trade has several defining characteristics, including (i) the need for foreign capital to locate in the importing country and establish a commercial presence, (ii) the frequent existence of market failures that offer a legitimate purpose for domestic regulations, and (iii) a lack of

---

of addressing the various problems in separable ways as I have just described, and in this sense is very different from the complications raised by "offshoring" and highlighted by Antras and Staiger (2012a, 2012b), where the nature of the pecuniary externality is itself altered. I return to this point in the Conclusion.

[26] The material in this section builds heavily on section 4 of Staiger and Sykes (2021).

readily available tariff-like instruments that can be applied to such trade, Staiger and Sykes establish that the key feature of mode 3 services trade that according to the terms-of-trade theory can account for the distinct design features of GATS relative to GATT is characteristic (iii), the lack of a readily available tariff-like instrument that can be applied to mode 3 services trade. Here I adopt this "missing tariff instrument" perspective on GATS and extend the partial equilibrium model of mode 3 services trade put forward by Staiger and Sykes to allow for the possibility of digital trade in services.

I focus on the policy choices of the home–country government, and to this end I assume that the service under consideration is demanded only in the home country. I represent this demand with the general downward-sloping demand curve $D(P)$ where D is a decreasing function and P is the consumer price of this service in the home-country market. The service must be produced in the home country, where it is consumed (mode 3). Moreover, to reflect the fact that market imperfections in the service sector are common, I assume that the consumption of this service generates a negative ("eyesore") externality that detracts from aggregate national welfare in the home country (the externality does not cross borders).[27] This externality is separate from any externality that might arise with open digital policies, and I will assume that it exists in both the pre-digital and digital world. To fix ideas the externality could be thought of as reflecting the noise or dust levels associated with construction services. The important role of this market imperfection is to introduce a legitimate public policy role for regulatory standards in the home-country service sector; the exact nature of the market imperfection is not crucial.

The home-country government can impose a regulatory standard s which specifies as a condition of entry into the home-country service market a (maximum) level of the externality $\phi(s)$ generated per unit of the service provided, and in principle a different standard may be applied to home-country and foreign service providers. I denote by r and ρ the particular standards imposed on home and foreign service providers respectively, with $\phi(r)$ and $\phi(\rho)$ the associated per-unit externality levels generated by consumption of the services provided by home and foreign service providers under their respective standards.[28] I assume that ϕ is

[27] As the service must be consumed where it is produced, it is immaterial whether the externality arises from the process of production or from the act of consuming the service: for ease of reference I assume it is consumption that generates the externality.

[28] My assumption here that the externality is incurred per-unit of the service demanded seems natural (e.g., for the noise levels associated with construction services), but differs from my

a decreasing and convex function. A nondiscriminatory standard would satisfy $r \geq \rho$, while $r < \rho$ would indicate that the home-country government has imposed a discriminatory standard on foreign service providers in the home-country market. Many policy interventions in the service sector take the form of entry restrictions, some implicit (licensing or certification requirements for entry) and some explicit (numerical quotas on numbers of service suppliers that may enter), and this modeling of regulatory standards can be thought of as a shorthand for such policy interventions.

To meet the standard s, service providers must incur a per-unit compliance cost, which includes both the cost of actually meeting the standard as well as the cost of establishing conformity with the standard. These costs, of course, are not immutable. And while it is natural that governments would be inclined toward design features of regulatory standards which favor local service providers and lead to higher costs of compliance for foreign service providers (see e.g., Copeland and Mattoo, 2008), it is also possible that government investment in the efficient design and implementation of a given standard could help to bring down these extra costs of compliance for foreign suppliers.

This possibility can be captured in a simple way as follows. I represent the compliance cost for home-country service firms facing standard s with the function $\kappa(s)$, where κ is increasing and convex in s. And I then assume that by investing $L \geq 0$ at a cost of $c_0 \cdot L$ in the design and implementation of the standard s, the home-country government can achieve a compliance cost for foreign service providers facing standard s equal to $\kappa^*(s, L) \equiv \kappa(s) + \lambda(L)$, where λ is decreasing and convex in L with $\lambda(L = 0) \equiv \bar{\lambda} > 0$ and $\lambda(\infty) \geq 0$. Hence, if the home-country government adopts the regulatory standard level s and chooses not to invest to reduce the cost of compliance for foreign service providers, then $\kappa^*(s, L = 0) = \kappa(s) + \bar{\lambda} > \kappa(s)$ and the cost of regulatory compliance to the standard level s in the home-country market will be greater for foreign than for home-country service providers. And by making investments in regulatory design, the cost of foreign compliance can be lowered toward the cost of home-country compliance. I assume that for any regulatory standards r and ρ, the quantity of services supplied by home-country and foreign service providers is then given respectively by the general upward-sloping supply curves $S_h(q_h - \kappa(r))$ for $q_h \geq \kappa(r)$ and $S_f(q_f - \kappa^*(\rho, L))$ for $q_f \geq \kappa^*(\rho, L)$, with S_h and S_f increasing functions and where

assumption about the form of the externality associated with digital openness, which I have assumed is independent of the amount of digital trade. As will become clear below, the findings I emphasize do not hinge on either of these assumptions.

q_h and q_f are the respective prices received in the home-country market by the home-country and foreign service providers.

I noted above that as a practical matter there is a "missing tariff instrument" when it comes to mode 3 services trade. At one level this is true by definition: import tariffs and export taxes are by definition not available in the context of mode 3 services trade, given that such trade does not cross international borders. Of course, in principle governments could have other fiscal instruments at their disposal that can mimic tariffs, such as a discriminatory sales tax imposed on foreign service providers at the point of sale of the mode 3 service. Following Staiger and Sykes (2021), I will ultimately assume that as a practical matter such policies are not available to governments due to the prohibitively high transaction costs that would be associated with their use. But for purposes of developing the benchmark model it is useful to abstract from these transaction-cost considerations and assume that such tariff-like instruments are available. For simplicity I will refer to the discriminatory sales tax imposed on foreign service providers as a "tariff," and denote the home-country and foreign tariffs by τ and τ^*, respectively, each expressed in specific terms. Finally, in addition to its standards policies and tariff, I assume that the home-country government can levy a nondiscriminatory sales tax t on home-country and foreign service providers alike (tax if positive and subsidy if negative), expressed in specific terms.[29]

As in the benchmark model of section 4, I assume that trade costs can be impacted by the digital policies of each country, which I again represent by the function $\iota(I, I^*)$, where $\iota(0, 0)$ is non-prohibitive and with $\iota(I, I^*)$ decreasing and convex in both its arguments and non-negative for all I and I^*. In the context of mode 3 services trade, I have in mind that these trade costs could include the cost of search to find a service provider (e.g., a web search versus a physical search), or the costs of delivery of the service (e.g., concentrating the logistics associated with the provision of the service in the foreign office or requiring these logistics to be performed on site in the home country); and as before, the digital policies I and I^* are taken to represent the kinds of regulatory barriers to digital trade described in section 3.[30]

[29] In light of the fact that any mode 3 service that is consumed in the domestic market must also be produced there, production taxes are equivalent to sales/consumption taxes in this setting. Therefore, while for simplicity I refer to these taxes as sales taxes, thinking of them as production taxes would be equally valid and leave unchanged the analysis to follow.

[30] For mode 3 services, where a commercial presence in the importing country must be established, it is still the case that the restrictive digital policies of each country could raise the cost of trade as I have assumed here. For example, data localization requirements can raise the

I restrict attention to taxes and trade costs that are set at non-prohibitive levels. Hence, the price paid by home-country consumers and the prices received by home-country and foreign service providers must satisfy

$$q_h + t = P = q_f + \iota(I, I^*) + \tau + \tau^* + t. \tag{5.1}$$

Note that services sell in the home country at the same price P regardless of the standard to which they are produced, reflecting the fact that individual consumers do not differentiate across units of the service on the basis of the magnitude of the externality ϕ that it generates when it is consumed.

The market-clearing condition $D(P) = S_h(q_h - \kappa(r)) + S_f(q_f - \kappa^*(\rho, L))$ determines equilibrium in this market. Using the pricing relationships in (5.1), this condition can be used to determine the market-clearing price received by foreign service providers

$$\hat{q}_f = \hat{q}_f(\overbrace{\iota(I, I^*) + \tau + \tau^*}^{(-)}, \overset{(-)}{t}, \overset{(+)}{r}, \overset{(+)}{\rho}, \overset{(-)}{L})$$

from which the market clearing price paid by home-market consumers and received by home-country service providers also follow:

$$\begin{aligned} \hat{P} &= \hat{q}_f(\iota(I, I^*) + \tau + \tau^*, t, r, \rho, L) + \iota(I, I^*) + \tau + \tau^* + t \\ &\equiv \hat{P}(\overbrace{\iota(I, I^*) + \tau + \tau^*}^{(+)}, \overset{(+)}{t}, \overset{(+)}{r}, \overset{(+)}{\rho}, \overset{(-)}{L}) \end{aligned}$$

$$\begin{aligned} \hat{q}_h &= \hat{q}_f(\iota(I, I^*) + \tau + \tau^*, t, r, \rho, L) + \iota(I, I^*) + \tau + \tau^*, \\ &\equiv \hat{q}_h(\overbrace{\iota(I, I^*) + \tau + \tau^*}^{(+)}, \overset{(-)}{t}, \overset{(+)}{r}, \overset{(+)}{\rho}, \overset{(-)}{L}). \end{aligned}$$

The derivative properties of the price functions are all intuitive and follow from the assumed properties of the demand and supply functions.

As I did in section 4, I can also define the *foreign world price* and the *home world price* of the mode 3 service by $P^{w*} \equiv q_f + \tau^*$ and $P^w \equiv P - \tau - t$, respectively. As before, the foreign and home world prices are related by $P^w - P^{w*} = \iota(I, I^*)$,

costs of trade by eliminating the option of cloud storage and thereby forcing a mode 3 service provider to invest in costly cyber security for its local data (see, for example, the discussion in Goldfarb and Trefler, 2019, and the reports listed in the Google Public Policy Blog posting https://publicpolicy.googleblog.com/2015/02/the-impacts-of-data-localization-on.html).

as can be confirmed with (5.1). A drop in trade costs ι brings P^w and P^{w*} closer together, and when $\iota = 0$ the home and foreign world prices are equated. The market-clearing home and foreign world prices are given by

$$\begin{aligned} \hat{P}^w &= \hat{q}_f(\iota(I,I^*)+\tau+\tau^*,t,r,\rho,L)+\iota(I,I^*)+\tau^* \\ &\equiv \hat{P}^w(\overbrace{\iota(I,I^*)+\tau^*}^{(+)},\overset{(-)}{\tau},\overset{(-)}{t},\overset{(+)}{r},\overset{(+)}{\rho},\overset{(-)}{L}) \end{aligned} \tag{5.2}$$

$$\begin{aligned} \hat{P}^{w*} &= \hat{q}_f(\iota(I,I^*)+\tau+\tau^*,t,r,\rho,L)+\tau^* \\ &\equiv \hat{P}^{w*}(\overbrace{\iota(I,I^*)+\tau}^{(-)},\overset{(+)}{\tau^*},\overset{(-)}{t},\overset{(+)}{r},\overset{(+)}{\rho},\overset{(-)}{L}). \end{aligned}$$

Following Staiger and Sykes (2011, 2021), it is also useful to define the market-clearing world prices of the foreign-produced unregulated ("raw") service – prior to bringing it into compliance with the prevailing home-country regulatory standard – and the associated foreign producer price of the unregulated service, by

$$\begin{aligned} \hat{P}_0^w &\equiv \hat{P}^w-\kappa^*(\rho,L)=\hat{P}_0^w(\overbrace{\iota(I,I^*)+\tau^*}^{(+)},\overset{(-)}{\tau},\overset{(-)}{t},\overset{(+)}{r},\overset{(-)}{\rho},\overset{(+)}{L}), \\ \hat{P}_0^{w*} &\equiv \hat{P}^{w*}-\kappa^*(\rho,L)=\hat{P}_0^{w*}(\overbrace{\iota(I,I^*)+\tau}^{(-)},\overset{(+)}{\tau^*},\overset{(-)}{t},\overset{(+)}{r},\overset{(-)}{\rho},\overset{(+)}{L}), \\ \hat{q}_f^0 &\equiv \hat{q}_f-\kappa^*(\rho,L)=\hat{q}_f^0(\overbrace{\iota(I,I^*)+\tau+\tau^*}^{(-)},\overset{(-)}{t},\overset{(+)}{r},\overset{(-)}{\rho},\overset{(+)}{L}). \end{aligned} \tag{5.3}$$

Notice that for any ρ and L there is a one-to-one correspondence between the world prices in (5.2) and the raw world prices in (5.3). But as Staiger and Sykes observe and as the top two lines of (5.3) confirm, it is the raw world prices that reveal the nature of terms-of-trade manipulation in this setting and which are most helpful for interpreting non-cooperative Nash policy choices. In particular, while both $\hat{P}^w$ and $\hat{P}_0^w$ are decreasing in τ and both $\hat{P}^{w*}$ and $\hat{P}_0^{w*}$ are increasing in τ^* reflecting the terms-of-trade improvement associated with an increase in one's own tariff as expected, and while both world price measures also exhibit the expected impact of trade costs ι, it is only the raw world prices $\hat{P}_0^w$ and $\hat{P}_0^{w*}$ that fall when the standard on foreign service providers ρ is raised, reflecting the ability of the home-country government to pass some of the cost of the higher standard on to foreign service providers, the essence of terms-of-trade manipulation in this regulatory setting. An analogous statement holds for the impact of L: as $\hat{P}_0^w$ and

$\hat{P}_0^{w*}$ reveal, when the home-country government invests more in lowering the cost of compliance for foreign service providers, foreign service providers capture part of the gains from this investment. For this reason I will follow Staiger and Sykes and use the raw world prices $\hat{P}_0^{w}$ and $\hat{P}_0^{w*}$ as the measures of each country's terms of trade.

Finally, as before I allow an open digital environment in a country to generate a non-pecuniary externality that takes an "eyesore" form. And as before, I assume that this externality may have both a local and a cross-border component, with the local component increasing in own-digital openness and the cross-border component increasing in the digital openness of one's trading partner. Home-country welfare is then given by the sum of consumer surplus for mode 3 services plus the producer surplus generated by home-country service providers plus tax revenue net of the cost of investment in the design and implementation of the standard, minus the cost of the eyesore externality generated by mode 3 service consumption and any externality associated with digital openness. And foreign-country welfare is the sum of the producer surplus generated by foreign-country service providers operating in the home-country market plus trade tax revenue, minus the cost of any externality associated with digital openness. And as before, it is straightforward to show that world welfare (the sum of home and foreign welfare) is independent of world prices.

5.2. The purpose of GATS in a pre-digital world

As I did with the benchmark trade-in-goods model, I will think of the pre-digital world as corresponding to a special case of the benchmark trade-in-services model described above in which

$$I \equiv 0 \equiv I^* \tag{5.4}$$

and there is no workable Internet in either country. As noted above, I will also assume that tariffs are unavailable to governments in the context of mode 3 services trade. But before imposing this instrument restriction, I first describe the efficient and Nash policy choices in the pre-digital world of mode 3 services when tariff-like instruments on mode 3 services trade are assumed to be available, as I assumed in developing the benchmark model above. In this setting, I demonstrate two points: first, I show that tariffs are not in fact needed to achieve efficiency; and second, I show that tariffs are the only policies that are distorted in the Nash equilibrium, just as was the case in the pre-digital world of the benchmark trade-in-goods model. With these results, I confirm a central finding of Staiger and

Sykes (2021) for the pre-digital world, namely, that the underlying purpose of a trade agreement is the same for both trade in goods and trade in mode 3 services. This is an important property of the pre-digital world, and in the next section I will ask whether it is preserved in the digital world.

The purpose of a trade agreement To proceed, I first record the policy choices that maximize world welfare in the pre-digital benchmark trade-in-services model:

$$\begin{aligned} \tau^e + \tau^{*e} &= 0 \\ t^e &= \phi(r^e) \\ [-\phi'(r^e)] - \kappa'(r^e) &= 0 = [-\phi'(\rho^e)] - \kappa'(\rho^e) \\ S_f^e \times [-\lambda'(L^e)] - c_0 &= 0. \end{aligned} \tag{5.5}$$

The interpretation of the efficient policies characterized in (5.5) is intuitive. From the third line of (5.5), the standards for both home-country and foreign service providers are chosen to equate the marginal benefit of a slightly higher standard in terms of per-unit externality reduction ($[-\phi'(r^e)]$ and $[-\phi'(\rho^e)]$) with the increase in marginal cost of service production from the slightly higher standard ($\kappa'(r^e)$ and $\kappa'(\rho^e)$): given the symmetric technologies across service providers in terms of both the externality that their services generate for a given standard ($\phi(s)$) and the sensitivity of the cost of compliance to changes in the standard ($\kappa(s)$), there is no efficiency reason to impose discriminatory standards across home-country and foreign service providers, and the third line of (5.5) implies $r^e = \rho^e$. With these efficient standards in place, the efficient sales tax (t^e) is then simply the Pigouvian tax that internalizes the remaining externality generated by the home-country and foreign service providers in the home-country market ($\phi(r^e)$), as the second line of (5.5) indicates. And it is efficient to invest in the design and implementation of the standard as it relates to foreign service providers to the point where the marginal benefit of such investment (the reduction in the total costs incurred by foreign service providers in meeting the standard, $S_f^e \times [-\lambda'(L^e)]$) is equal to the marginal cost of the investment (c_0), as the bottom line of (5.5) reflects. Notice that while the efficient standard applied to both home-country and foreign service providers is the same, the efficient investment in reducing the cost of compliance for foreign service providers leaves these providers with a higher regulatory cost than that faced by home-country service providers. Finally, as indicated by the top line of (5.5), with these efficient interventions in place there is no role for tariffs in the set of efficient policy interventions.

I next turn to the noncooperative Nash policies that each government would choose unilaterally to maximize its own welfare. Straightforward calculations reveal that the Nash policies are given by:

$$\begin{aligned} \tau^N &= \frac{\hat{P}^{w*N}}{\eta^{E*N}} \text{ and } \tau^{*N} = \frac{\hat{P}^{wN}}{\eta^{M^N}} \\ t^N &= \phi(r^N) \\ [-\phi'(r^N)] - \kappa'(r^N) &= 0 = [-\phi'(\rho^N)] - \kappa'(\rho^N) \\ S_f^N \times [-\lambda'(L^N)] - c_0 &= 0, \end{aligned} \tag{5.6}$$

where I use the notation $\eta^{E^{*N}}$ to refer to the elasticity of foreign export supply (i.e., the elasticity of mode 3 supply of foreign service providers in the home-country market) evaluated at Nash policies, and I use the notation η^{M^N} to refer to the elasticity of home import demand defined positively (i.e., the elasticity of home-country demand for mode 3 services supplied by foreign service providers, defined positively) evaluated at Nash policies. Evidently, as a comparison of (5.5) and (5.6) confirms, only the tariffs on mode 3 services trade are inefficient in the Nash equilibrium of the pre-digital world, and they take the familiar Johnson optimal tariff form: conditional on Nash trade volumes, all other Nash policy choices are efficient. This confirms a central finding of Staiger and Sykes (2021) for the pre-digital world, namely, that despite a number of salient differences between goods and services, the underlying purpose of a trade agreement in the context of mode 3 services trade is the same as the purpose of a trade agreement over goods trade: in both cases, that purpose is to reduce tariffs and expand market access and trade volumes to efficient levels.

The design of GATS Having established for the pre-digital world that the purpose of a trade agreement is the same across goods and (mode 3) services, the design of GATS seems puzzling. After all, if (5.6) describes the starting point from which governments considered the design of a trade-in-services agreement, the strategy of borrowing heavily from the design features of GATT – and therefore focusing market access negotiations on a single policy instrument while putting in place various rules to prevent governments from back-sliding on the market access commitments implied by their negotiated bindings with new protective behind-the-border measures – seems like an obvious and natural way to proceed. What, then, explains the striking difference between GATT and GATS that I noted at the outset of this section? I now illustrate the key elements of the "missing tariff instrument" interpretation of this difference put forward by Staiger and Sykes (2021).

To this end, I introduce the following assumption, eliminating each government's ability to impose tariffs on mode 3 services trade:

$$\textit{No tariffs on mode 3 services trade: } \tau = \tau^* \equiv 0. \qquad \text{(Assumption 2)}$$

Assumption 2 can be rationalized in practical terms on the grounds of the prohibitively high transaction costs that would be associated with the policy instruments τ and τ^* given that in the context of mode 3 services trade they represent discriminatory sales taxes on foreign mode 3 service providers administrated at the point of sale.

As (5.5) confirms, Assumption 2 does not alter the efficient setting of non-tariff policies in this environment, because tariffs are not part of the efficient policy package. But the absence of a tariff for purposes of terms-of-trade manipulation has a profound impact on the unilateral (Nash) non-tariff policy choices of the home-country government, which are now given by the following:

$$\begin{aligned}
t^N - \phi(r^N) &= \left[\frac{\Theta^N}{S_h' + S_f'}\right] > 0 \\
[-\phi'(r^N)] - \kappa'(r^N) &= \left[\frac{\Theta^N}{S_h' + S_f'}\right] \times \left[\frac{S_h' \times \kappa'(r^N)}{S_h^N}\right] > 0 \\
[-\phi'(\rho^N)] - \kappa'(\rho^N) &= \left[\frac{-\Theta^N}{S_h' + S_f'}\right] \times \left[\frac{S_h' \times \kappa'(\rho^N)}{S_f^N}\right] < 0 \\
S_f^N \times [-\lambda'(L^N)] - c_0 &= \left[\frac{\Theta^N}{S_h' + S_f'}\right] \times S_h' \times [-\lambda'(L^N)] > 0,
\end{aligned} \tag{5.7}$$

where $\Theta^N \equiv (S_f^N - S_f' \times [\phi(r^N) - \phi(\rho^N)])$, and where under the unilaterally optimal home-country policies $\Theta^N > 0$. As a comparison of (5.7) with (5.6) confirms, without its tariff-equivalent policy instrument, the home-country government must turn to its other policies as second-best means to manipulate the terms of trade. And as a comparison of (5.7) with (5.5) reveals, to this end the government will set its nondiscriminatory sales tax above the Pigouvian level ($t^N > \phi(r^N)$), impose a lower-than-efficient standard on home-country service providers ($[-\phi'(r^N)] > \kappa'(r^N)$) and a higher-than-efficient standard on foreign service providers ($[-\phi'(\rho^N)] < \kappa'(\rho^N)$), and make smaller-than-efficient compliance-cost-reducing investments in the design and implementation of the standard applied to foreign service providers ($S_f^N \times [-\lambda'(L^N)] > c_0$). Each of these policy distortions can be understood from examining their terms-of-trade implications as recorded in (5.3).

In short, and as Staiger and Sykes (2021) argue, without a tariff to manipulate its services terms of trade, an importing government will in the Nash equilibrium tend to spread protective distortions widely across the policy instruments that it does wield in the service sector, thereby “contaminating” many of its Nash policies with internationally inefficient terms-of-trade motives. And if (5.7) describes the starting point from which governments would have considered the design of a trade-in-services agreement, the strategy of borrowing heavily from the shallow-integration design features of GATT no longer seems like an obvious and natural, or even viable, way to proceed, so much so that it seems plausible that this strategy may not have even occurred to GATS negotiators. Rather, with (5.7) as their starting point and facing evident behind-the-border policy distortions spread throughout the domestic service market, a decision to adopt a deep-integration approach to services liberalization seems almost inevitable. From this perspective the lack of an effective tariff or tariff-equivalent policy instrument for (mode 3) service-sector intervention can go a long way in accounting for the striking differences in the architecture of GATS and GATT.

A GATS redesign for the pre-digital world Nevertheless, as Staiger and Sykes (2021) demonstrate, this does not mean that shallow integration is impossible for services. Rather, Staiger and Sykes show that it is possible to exploit the elements of the trade-in-goods problem that are shared with the trade-in-services problem to devise a “two-step” path forward for liberalizing trade in services that has much in common with the shallow-integration approach of GATT.

As Staiger and Sykes (2021) describe, in a first step governments would agree to a set of blanket rules to apply to services along the lines of the GATT rules that apply to goods, namely: (i) the national treatment (NT) rule, which prohibits domestic regulatory and tax policies that discriminate against foreign trade; (ii) the agreement on technical barriers to trade (TBT), which prohibits unnecessarily trade restrictive regulatory choices; and (iii) the non-violation (NV) clause, which protects the value of market access concessions from erosion due to subsequent and unanticipated changes in non-contracted policies. Staiger and Sykes show that, in ruling out discriminatory and unnecessarily trade restrictive regulatory choices, such an agreement would induce governments to unilaterally remove protectionist elements from their standards and regulatory policies in the service sector and divert them into a narrow set of fiscal – but not regulatory – measures, or perhaps into a single preferred regulatory measure. And with international policy inefficiencies concentrated in such a limited set of instruments, governments could then in a second step use negotiations over these instruments to establish (in concert with the NT, TBT and NV rules) the market access commitments in service sectors needed to arrive at the efficiency frontier, without the need to directly negotiate over a wide range of domestic regulatory measures, much as GATT has used negotiated commitments on tariffs in the goods sector. As Staiger and Sykes argue, the possibility of pursuing shallow integration of services trade in this way might allow countries to side-step the difficult sovereignty issues that arise with deep integration efforts, and that may help explain the relative lack of GATS's success to date as compared to GATT.

As should be clear from the discussion above, the property of the pre-digital world that underpins the possibility of a shallow-integration approach to liberalization in the service sector is that the problem for a trade-in-services agreement to solve is not so different from the problem that a trade-in-goods agreement must solve. And the key step in establishing this property is to show that, absent the missing tariff for mode 3 services – that is, when Assumption 2 is not imposed – the two problems are the same. In the next section I consider whether this property extends to a digital world.

5.3. The purpose of GATS in a digital world

In a world of digital trade in services, the benchmark model of section 5.1 applies. As I did when considering the pre-digital world, below I first characterize efficient policies, then characterize Nash policies, and then consider the purpose of a trade agreement covering services trade in a digital world. Again I proceed by considering two possibilities for the digital world, one where the non-pecuniary externality associated with digital openness is purely local and does not spill over across borders, and the other where this externality has both local and cross-border components.

5.3.1. A purely local non-pecuniary externality from digital openness

I consider first the case of a purely local non-pecuniary externality from digital openness, by imposing Assumption 1. For the reasons explained just above, in what follows I assume that countries have access to tariffs on (mode 3) digital services trade, and therefore I do not impose Assumption 2.

It is direct to calculate that for this case the efficient policies when countries engage in digital trade in services are characterized by

$$\begin{aligned} \tau^e + \tau^{*e} &= 0 \\ t^e &= \phi(r^e) \\ [-\phi'(r^e)] - \kappa'(r^e) &= 0 = [-\phi'(\rho^e)] - \kappa'(\rho^e) \\ S_f^e \times [-\lambda'(L^e)] - c_0 &= 0 \\ S_f^e \times [-\frac{\partial \iota}{\partial I}] - c'(I^e) &= 0 = S_f^e \times [-\frac{\partial \iota}{\partial I^*}] - c^{*\prime}(I^{*e}). \end{aligned} \tag{5.8}$$

Comparing (5.8) with (5.5), the only difference in efficient trade-in-services policies between the pre-digital and digital world involves the efficient setting of the digital policies I and I^*, which are absent from (5.5) and characterized in the last line (5.8). The interpretation of the efficient digital policies I^e and I^{*e} for the case of digital trade in services and a purely local non-pecuniary externality is analogous to the interpretation for the case of digital trade in goods.

What about the Nash policies for this case? The Nash policies when countries engage in digital trade in services are characterized by

$$\begin{aligned} \tau^N &= \frac{\hat{P}^{w*N}}{\eta^{E*N}} \quad \text{and} \quad \tau^{*N} = \frac{\hat{P}^{wN}}{\eta^{MN}} \\ t^N &= \phi(r^N) \\ [-\phi'(r^N)] - \kappa'(r^N) &= 0 = [-\phi'(\rho^N)] - \kappa'(\rho^N) \\ S_f^N \times [-\lambda'(L^N)] - c_0 &= 0 \\ S_f^N \times [-\frac{\partial \iota}{\partial I}] - c'(I^N) &= 0 = S_f^N \times [-\frac{\partial \iota}{\partial I^*}] - c^{*\prime}(I^{*N}). \end{aligned} \tag{5.9}$$

But as a comparison of (5.9) with (5.8) reveals, when countries have access to a tariff-like policy instrument to apply to (mode 3) digital services trade and when the non-pecuniary externality from digital openness is purely local, only the tariffs on services trade are inefficient in the Nash equilibrium of the digital world, and they take the familiar Johnson optimal tariff form: conditional on Nash trade volumes, all other Nash policy choices are efficient, just as in the pre-digital world.

I have therefore demonstrated with (5.8) and (5.9) that the key step in establishing the possibility of a shallow-integration approach to liberalization in the service sector in the pre-digital world extends to a digital world for the case of a purely local non-pecuniary externality from digital openness; and from here it is straightforward to show that the rest of the argument extends as well. Evidently, then, in the digital world the possibility of a shallow-integration approach to liberalization in the service sector remains, just as Staiger and Sykes (2021) describe that possibility for the pre-digital world.

5.3.2. A cross-border non-pecuniary externality from digital openness

When the non-pecuniary externality from digital openness flows across borders, Assumption 1, which I imposed in the previous section, no longer holds. And as with the analysis of digital trade in goods from section 4, when Assumption 1 is violated and $\theta > 0$, the Nash choices of digital openness in the presence of digital services trade are unaffected but the efficient choices are now characterized by

$$S_f^e \times [-\frac{\partial \iota}{\partial I}] = [1+\theta] \times c'(I^e) \tag{5.10}$$

for the home-country government and by

$$S_f^e \times [-\frac{\partial \iota}{\partial I^*}] = [1+\theta] \times c^{*\prime}(I^{*e}) \tag{5.11}$$

for the foreign-country government.

With (5.10) and (5.11) exactly analogous to (4.18) and (4.19), it is clear that all of the conclusions reached in section 4 for the case of digital trade in goods in the presence of a cross-border non-pecuniary externality from digital openness carry over to the case of digital trade in services. In particular, as I noted in section 4, in the presence of a cross-border non-pecuniary externality from digital openness there are now two problems to solve: the cross-border non-pecuniary externality must be addressed; and conditional on addressing this first problem, the insufficient market access problem familiar from the pre-digital world must also be addressed. But as I also noted in section 4, even in this case there may be an approach to integration that lies somewhere between shallow integration on the one hand, and on the other hand fully deep integration where countries negotiate over all aspects of digital policy that enter into I and I^*.

6. Digital Trade and the Design of the WTO

In this section I assess what the analysis presented in the previous sections means for the design of the WTO in a world of digital trade. The main message comes in two parts.

Purely local non-pecuniary externalities First, if the non-pecuniary externalities associated with open digital policies are purely local, then the purpose of a trade agreement in a digitalized world is to achieve an efficient level of market access expansion, the same purpose as in the pre-digital world. And for this purpose a shallow-integration approach can in principle suffice.[31]

Of course, in the analysis above I have abstracted from a number of challenges that a shallow approach to integration would confront in practice;[32] and in reality the line between shallow and deep integration is not as stark as my analysis makes it out to be. At a more practical level, therefore, and as I emphasized earlier, the message in this case is not so much that *no* degree of deep integration is necessary in a digital world, but rather that the existing market access orientation of the WTO can provide a potentially useful *guardrail* to delineate the "depth" of integration that trade agreements should contemplate in the digital world: according to the findings above, and just as in the pre-digital world, there is no reason for a trade agreement in the world of digital trade to go deeper than what is required to ensure that property rights over negotiated market access are reasonably secure. The use of such a guardrail could help governments avoid conflicts between globalization and national sovereignty that, according to the analysis presented above, would be unnecessary.

Even in this case, the analysis above suggests that the rise of digital trade could pose a challenge for the WTO, because it has the potential to disrupt the market access implications of existing WTO commitments. For example, while a good delivered via instructions for 3-D printing is still a good according to my taxonomy, it may be that as a practical matter the tariff commitments negotiated in the pre-digital world no longer afford the same degree of trade protection against imports of that good that they once did; and it may also be that new digital forms of protection become relevant which could undermine existing market access commitments. But at least in principle, existing WTO rules are well designed to handle these issues, because these issues are not new to the digital age. The first

[31] Interestingly, adopting a very different modeling approach that features monopoly platforms and two-sided markets for platform services and focuses on the particular issue of privacy protection, McCalman (2021) comes to a similar conclusion that shallow integration may suffice for issues related to digital trade.

[32] See Staiger and Sykes (2021) for a discussion of a number of these challenges in the context of a shallow-integration approach to the liberalization of trade in services.

of these issues could be handled under the various renegotiation clauses that exist in the WTO, which are designed to allow WTO market access commitments to function as "liability rules" that can be renegotiated without fear of hold up from trading partners.[33] And the second issue can in principle be addressed with the existing WTO rules (e.g., NT, TBT, NV) that are designed to handle such issues more broadly.

There is also an interesting further implication of the analysis above in the case where the non-pecuniary externalities associated with open digital policies are purely local: the existing moratorium on tariffs on electronic transmissions might be *complicating* the task of shallow integration in a world of digital trade. This is because, as the analysis above makes clear, it is only when constraints on the use of tariffs on digital trade are introduced that the unilateral choices of non-tariff behind-the-border policies become distorted for protectionist terms-of-trade-manipulation purposes. Viewed from this perspective, in the case where the non-pecuniary externalities associated with open digital policies are purely local, the suspension or termination of the moratorium on tariffs on electronic transmissions should have the effect of inducing governments unilaterally to concentrate their protective and internationally inefficient measures on such trade into tariffs; and in inducing such "tariffication," abandoning this moratorium might represent a useful first step toward effective shallow integration in a digital world.

A final pair of observations that apply to this first case are also relevant. First, even if the purpose of a trade agreement does not change with the rise of digital trade, the blurring of the distinction between goods and services that digitalization has caused and will increasingly cause in the future presents an important challenge for the WTO in light of the dramatically different structure of GATT and GATS. This challenge can be reduced if the design features of GATT and GATS are brought closer together, since then the distinction between goods and services becomes less consequential within the WTO framework. As I have described above, in principle GATS could be redesigned to be more like GATT. The advent of the digital age may make such a redesign all the more attractive.

At the same time, the analysis of the previous sections also suggests a second, related observation: from a functional perspective, a new approach to the classifi-

[33] In the context of the WTO, renegotiations of market access commitments for goods are handled through GATT Article XXVIII, while for services they are handled through GATS Article XXI (see Hoda, 2001, on the history of renegotiations in GATT and the early WTO). On the liability rule interpretation of WTO market access commitments, see Schwartz and Sykes (2002), Pauwelyn (2008) and Maggi and Staiger (2015).

cation of goods and services for the digital world might be attractive. Recall that for the analysis above I have adopted a simple taxonomy that partitions digital trade transactions into two groups: digital trade in goods, and digital trade in services; and to develop this taxonomy I have defined a transaction as involving a good (service) if at the moment of consumption that transaction is a good (service) as traditionally defined. And as I noted, for the formal analysis I have assumed that digitalization does not itself alter the feasibility of imposing a tariff on the trade in question. My analysis, however, has confirmed for the digital world what the analysis of Staiger and Sykes (2021) implied for the pre-digital world, namely, that for the purposes of the design features of a trade agreement what functionally distinguishes trade in goods from trade in services is whether or not a tariff can be feasibly applied to that trade. This suggests the attractiveness of adopting a simple alternative classification system for traded goods and services within the context of the WTO: simply put, digital or otherwise, traded goods would refer to transactions on which a tariff can feasibly be applied, and these transactions would be covered under GATT; traded services would refer to transactions on which a tariff cannot be applied, and these transactions would be governed under GATS.[34] Under this alternative classification, digitalization would change a good into a service or vice versa if and only if it altered the feasibility of imposing tariffs on the international transaction.

Cross-border non-pecuniary externalities The second part of the message applies where the non-pecuniary externalities associated with open digital policies cross international borders. As I have noted, in this case the purpose of a trade agreement becomes more complex, as there are now two problems to solve: the cross-border non-pecuniary externality must be addressed; and conditional on addressing this first problem, the insufficient market access problem familiar from the pre-digital world must also be addressed. But I also emphasized a further point: even in this case, there may be an approach to integration for goods and services trade in a digital world that lies somewhere between GATT's shallow integration approach on the one hand, and on the other hand fully deep integration

[34] This alternative classification adopts a functional perspective in the sense that it is based on the answer to the question "Can a tariff be applied to the transaction?"; see Willemyns (2019) for a disussion of alternative functional and other approaches to the classification of services in the WTO. It should also be noted that the alternative classification I describe here would not necessarily conform to the principal of "technological neutrality" that currently applies to GATS commitments (see the WTO Analytical Index Annex 1B at https://www.wto.org/english/res_e/publications_e/ai17_e/gats_art1_jur.pdf.

where countries negotiate over all aspects of digital policy. This is because if countries could jointly focus on just those aspects of their digital openness that are generating the cross–border non-pecuniary externalities and address those externalities with limited negotiations, then, from that point forward, we would be back in a world where there are no (unaddressed) cross-border non-pecuniary externalities and where only local non-pecuniary externalities remain; and as the analysis above confirms, that is a world where achieving the efficiency frontier with GATT's shallow approach to integration is possible.

To illustrate this point, I now briefly describe the possibility of a middle ground between shallow and deep integration in the context of the "data de-correlation" scheme proposed by Acemoglu et al (forthcoming). Acemoglu et al consider a situation in which users of a digital platform value privacy and impose negative externalities on each other when they share their personal data with the platform, provided that their data is correlated with other users so that the platform also learns something about other users through this correlation. In such a situation, individual-level data is underpriced and the market economy generates too much data. As Acemoglu et al describe, data de-correlation represents one possible solution to address this problem. De-correlation is a scheme for mediating (via a trusted third party) data transactions in a way that reduces their correlation with the data of other users – and in particular the correlation between the data of a user who is not sharing her data with the data of others who have shared their data – and thereby mitigates these externality-induced privacy concerns.[35]

In the context of the present discussion, we could think of WTO member governments agreeing to a limited form of this proposal, tailored to address just the correlation with the data of other *international* users – and hence addressing just the cross-border non-pecuniary externality associated with these digital privacy issues – and leaving the handling of correlation of users' data *within* national borders to the discretion of each national government. In this way, users of the digital platform would retain private property rights over their own cross-border data flows and thereby avoid the associated cross-border non-pecuniary externality that would otherwise arise, and such property rights might be protected within the WTO Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS); but whether or not users retained such rights over their own within-

[35] As Acemoglu et al (forthcoming) discuss, the de-correlation scheme that they propose is different from procedures which anonymize data, because anonymization hides information about the user who is sharing her data, while de-correlation only hides information about others who are correlated with this user.

country data flows would be a decision for each national government to make.

7. Conclusion

The design of a trade agreement should reflect its purpose. Does digital trade change the purpose of a trade agreement? To explore this question in a formal model, I have adopted a simple taxonomy of the ways in which digital trade can arise and the policies that can be used to impact such trade. I have reviewed what the theoretical literature on the economics of trade agreements has to say about the purpose of a trade agreement covering both goods and services in a pre-digital model world economy, and how this purpose can be seen to be reflected in the broad design features of both GATT and GATS, the WTO agreements that govern international trade in goods and services respectively. I have then introduced digital trade into the model world economy and revisited the purpose of a trade agreement; and from this perspective I have considered whether the rise of digital trade warrants changes in the design of the WTO.

With digital policies, as with other behind-the-border measures, the fact that regional and mega-regional agreements are delving deeply into efforts to negotiate directly over such measures does not necessarily mean that the WTO should follow suit, even if it could. Indeed my findings suggest that the WTO's shallow approach to integration is perhaps better designed to deal with digital trade than is commonly believed. I have shown that where non-pecuniary externalities associated with digital openness exist but are purely local, the purpose of a trade agreement for both goods trade and trade in services is unchanged by the advent of the digital world. I have demonstrated that this implies that, in such a world, the existing shallow-integration features that worked well for GATT in the pre-digital world can in principle be applied to digital policies impacting goods trade; and I have argued that a GATT-like shallow-integration approach to trade in services which departs from the approach taken by GATS but is along the lines proposed by Staiger and Sykes (2021) could in principle be applied to digital policies impacting services trade as well. I have also shown that where the non-pecuniary externalities associated with digital openness cross international borders, the purpose of a trade agreement becomes more complex. But I have demonstrated that even in this case there may be an approach to integration for goods and services trade in a digital world that lies somewhere between GATT's shallow integration approach on the one hand, and on the other hand fully deep integration where countries negotiate over all aspects of digital policy.

I have taken a reduced-form approach to modeling the impact of digitalization on the world economy, and in doing so I have abstracted from many important features of digital trade and digital policies. If modeled, a number of these features would introduce potentially important departures from the simple competitive industry environment that I have adopted for my formal analysis, and it is possible that such departures could alter my findings. As I noted earlier, the terms-of-trade theory of trade agreements has been shown to yield robust findings across a wide range of market structures, and I suspect that my findings here will prove fairly robust as well. Still, it is important to move beyond the reduced-form approach I have adopted and investigate the impacts of digitalization on the purpose of a trade agreement in micro-founded models of digital trade. In this regard, Antras and Staiger (2012a, 2012b) have shown that the terms-of-trade theory does not extend to "offshoring" settings with the characteristic that domestic and foreign buyers exchange specialized inputs whose prices are determined by bilateral bargaining and are not disciplined by industry-wide market clearing conditions.[36] To the extent that digital trade, and especially digital trade in services along a global value chain, is thought to exhibit this characteristic, an important caveat to the results I have derived here should be kept in mind. Extending the analysis to such an environment is an important direction for future work.

References

[1] Aaronson, Susan A. 2018. "What are we talking about when we talk about digital protectionism?" *World Trade Review*, August: 1-37.

[2] Acemoglu, Daron, Makhdoumi, Ali, Malekian, Azarakhsh and Asu Ozdaglar. Forthcoming. "Too Much Data: Prices and Inefficiencies in Data Markets," *American Economic Journal: Microeconomics*.

[3] Ahmed, Usman. 2019. "The Importance of Cross-Border Regulatory Cooperation in an Era of Digital Trade." *World Trade Review* 18(S1): s99-s120.

[4] Allen, Treb. 2014. "Information Frictions in Trade," *Econometrica* 82(6): 2041-2083.

[36] See also Bagwell, Bown and Staiger (2016) and Grossman, McCalman and Staiger (2021) for other modeling frameworks where the shallow-integration results of the terms-of-trade theory need not hold.

[5] Antras, Pol and Robert W. Staiger. 2012a. "Offshoring and the Role of Trade Agreements," *American Economic Review*, 102(7): 3140-3183.

[6] Antras, Pol and Robert W. Staiger. 2012b. "Trade Agreements and the Nature of Price Determination," *American Economic Review Papers and Proceedings*, 102(3): 470-476.

[7] Azmeh, Shamel, Foster, Christopher and Jaime Echavarri. 2020. "The International Trade Regime and the Quest for Free Digital Trade," *International Studies Review: Analytical Essay* 22: 671-692.

[8] Bacchetta, Marc, Low, Patrick, Mattoo, Aaditya, Schuknecht, Ludger, Wagner, Hannu and Madelon Wehrens. 1998. *Electronic Commerce and the Role of the WTO*, World Trade Organization Special Studies 2.

[9] Bagwell, Kyle, Bown, Chad, and Robert W. Staiger. 2016. "Is the WTO Passe?," *Journal of Economic Literature* 54(4): 1125-1231.

[10] Bagwell, Kyle and Seung Hoon Lee. 2020. "Trade policy under monopolistic competition with firm selection," *Journal of International Economics* 127: 1-39.

[11] Bagwell, Kyle and Robert W. Staiger. 1999. "An Economic Theory of GATT," *American Economic Review* 89(1): 215-248.

[12] Bagwell, Kyle and Robert W. Staiger. 2001. "Domestic Policies, National Sovereignty, and International Economic Institutions," *Quarterly Journal of Economics* 116: 519-562.

[13] Bagwell, Kyle and Robert W. Staiger. 2002. *The Economics of the World Trading System*, Cambridge: The MIT Press.

[14] Bagwell, Kyle and Robert W. Staiger. 2016. "The Design of Trade Agreements," in Kyle Bagwell and Robert W. Staiger (eds) *The Handbook of Commercial Policy*, vol 1A, December.

[15] Blum, Bernardo S. and Avi Goldfarb. 2006. "Does the internet defy the law of gravity?" *Journal of International Economics* 70: 384-405.

[16] Broda, Christian, Nuno Limão, and David E. Weinstein. 2008. "Optimal tariffs and market power: the evidence." *American Economic Review* 98(5): 2032-65.

[17] Casalini, F., J. Lopez Gonzalez and E. Moise. 2019. "Approaches to market openness in the digital age", *OECD Trade Policy Papers*, No. 219, OECD Publishing, Paris. http://dx.doi.org/10.1787/818a7498-en.

[18] Castro, Daniel and Alan McQuinn. 2015. *Cross-border data flows enable growth in all industries*, Information Technology and Innovation Foundation,February, http://www2.itif.org/2015-cross-border-data-flows.pdf.

[19] Ciuriak, Dan. 2019. "World Trade Organization 2.0: Reforming Multilateral Trade Rules for the Digital Age," *Policy Brief No. 152*, Centre for International Governance Innovation, July.

[20] Copeland, Brian and Aaditya Mattoo. 2008. "The Basic Economics of Services Trade." in A. Mattoo, R. M. Stern and G. Zanini (eds) *A Handbook of International Trade in Services*, Oxford University Press.

[21] Disdier, Anne-Celia and Keith Head. 2008. "The puzzling persistence of the distance effect on bilateral trade," *Review of Economics and Statistics* 90(1): 37–48.

[22] Freund, Caroline L. and Diana Weinhold. 2004. "The effect of the Internet on international trade," *Journal of International Economics* 62: 171-189.

[23] Friedman, Thomas L. 2005. *The World is Flat: A brief history of the twenty-first century*, Farrar, Straus and Giroux, New York.

[24] Goldfarb Avi and Catherine Tucker. 2019. "Digital Economics," *Journal of Economic Literature* 57(1), March: 3–43.

[25] Goldfarb, Avi and Daniel Trefler. 2019. "Artificial Intelligence and International Trade," in Ajay Agrawal, Joshua Gans and Avi Goldfarb (eds.) *The Economics of Artificial Intelligence: An Agenda*, National Bureau of Economic Analysis, Cambridge, MA: 463–92.

[26] Grossman, Gene M., McCalman, Phillip and Robert W. Staiger. 2021. "The 'New' Economics of Trade Agreements: From Trade Liberalization to Regulatory Convergence," *Econometrica* 89(1): 215-249.

[27] Head, Keith and Thierry Mayer. (2014). "Gravity equations: workhorse, toolkit, and cookbook" in Kenneth Rogoff, Elhanan Helpman and Gita

Gopinath (eds.) *Handbook of International Economics*, vol. 4, pp. 131–95, Amsterdam: Elsevier.

[28] Hoda, Anwarul. 2001. *Tariff Negotiations and Renegotiations under the GATT and the WTO: Procedures and Practices*, WTO and Cambridge University Press.

[29] Johnson, Harry G. 1953-4. "Optimum Tariffs and Retaliation," *Review of Economic Studies* 21(2): 142–53.

[30] Leamer, Edward E. 2007. 'A flat world, a level playing field, a small world after all, or none of the above? A review of Thomas Friedman's the world is flat', *Journal of Economic Literature* 45(1): 83–126.

[31] Lendle, Andreas, Marcelo Olarreaga, Simon Schropp, and Pierre-Louis Vézina. 2016. "There goes gravity: eBay and the death of distance," *The Economic Journal* 126: 406-441.

[32] Maggi, Giovanni and Robert W. Staiger. 2015. "Optimal Design of Trade Agreements in the Presence of Renegotiation," *American Economic Journal: Microeconomics* 7(1): 109-143.

[33] McCalman, Phillip. 2021. "e-Globalization and Trade Agreements," University of Melbourne, mimeo, May.

[34] Meltzer, Joshua P. 2013. "The Internet, Cross-Border Data Flows and International Trade," *Issues in Technology Innovation* 22, February.

[35] Meltzer, Joshua P. 2019. "Governing Digital Trade," *World Trade Review* 18(S1): s23-s48.

[36] Pauwelyn, Joost. 2008. *Optimal Protection of International Law: Navigating between European Absolutism and American Volunteerism.* Cambridge: Cambridge University Press.

[37] Petersmann, Ernst-Ulrich. 1997. *The GATT/WTO Dispute Settlement System: International Law, International Organization and Dispute Settlement.* London: Kluwer Law International.

[38] Schwartz, Warren F. and Alan O. Sykes. 2002. "The Economic Structure of Renegotiation and Dispute Resolution in the World Trade Organization." *Journal of Legal Studies* 31(1): S179–204.

[39] Smith, Fiona, and Lorna Woods. 2005. "A Distinction Without a Difference: Exploring the Boundary between Goods and Services in the World Trade Organization and the European Union." *Yearbook of European Law* 24(1): 463–510.

[40] Staiger, Robert W. 2018. "On the implications of digital technologies for the multilateral trading system," *WTO World Trade Report*, Opinion Piece: 150.

[41] Staiger, Robert W. 2019. "Non-Tariff Measures and the WTO," in Ben Zissimos (ed.), *The WTO and Economic Development*, MIT Press.

[42] Staiger, Robert W. Forthcoming. *A World Trading System for the Twenty-First Century*, MIT Press.

[43] Staiger, Robert W. and Alan O. Sykes. 2011. "International Trade, National Treatment, and Domestic Regulation," *Journal of Legal Studies* 40(1): 149-203.

[44] Staiger, Robert W. and Alan O. Sykes. 2021. "The Economic Structure of Trade-in-Services Agreements," *Journal of Political Economy* 129(4): 1287-1317.

[45] UNCTAD. 2019. *Digital Economy Report 2019*, United Nations Conference on Trade and Development, https://unctad.org/system/files/official-document/der2019_en.pdf.

[46] USITC. 2013. *Digital Trade in the U.S. and Global Economies, Part 1*, Investigation No. 332-531, USITC Publication 4415, July.

[47] USITC. 2014. *Digital Trade in the U.S. and Global Economies, Part 2*, Investigation No. 332-540, USITC Publication 4485, August.

[48] Willemyns, Ines. 2019. "GATS Classification of Digital Services – Does 'The Cloud' Have a Silver Lining?," *Journal of World Trade* 53(1): 59–82.

[49] WTO. 1998. *Work Programme on Electronic Commerce.* Geneva: World Trade Organization.

[50] WTO. 2018. *World Trade Report. The future of world trade: How digital technologies are transforming global commerce.* Geneva, World Trade Organization.

[51] Wu, Mark. 2017. "Digital Trade-Related Provisions in Regional Trade Agreements: Existing Models and Lessons for the Multilateral Trade System," *RTA Exchange Overview Paper*, Inter-American Development Bank, International Centre for Trade and Sustainable Development, November.