What does Regional Economic Integration Deliver?*

Naoto Jinji†

Faculty of Economics, Kyoto University

Abstract
I examine whether recent regional integrations are deeper than older ones using data on the contents of trade agreements. I show that although recent regional integrations are deeper than older ones, the coverage of policy areas is mainly restricted to the provisions that fall under the current mandate of the WTO discipline. I also show that the depth and breadth of RTAs are heterogeneous according to types of RTAs, country pairs, and particular signatory countries, though the differences tend to be narrowing. Then, reviewing the findings of recent studies, I argue that deep integration delivers many gains to trade in goods, production networks, and research and development. However, further studies are required to fully understand the gains and losses of deep integration.

Key words: regional economic integration; shallow and deep integration; regional trade agreements.

JEL classification: F13; F14; F15

1 Introduction
The year 2016 may be marked as a turning point of globalization. On June 24, 2016, the people of the United Kingdom voted for Brexit in a referendum. On November 8, 2016, Donald Trump, who proposed the “America First” policy as well as various protectionism policies including the construction of a substantial wall on the United States–Mexico border during his presidential campaign, won the US presidential election. These two events are symbolic of the curbing of the trend of globalization. After the Brexit referendum, the United Kingdom struggled to decide whether and when it would actually leave the European Union (EU) for more than three years. Finally, it withdrew from the EU on January 31, 2020. In the United States, the Trump administration implemented protectionism policy. In particular, it increased tariffs...
on imports from China, causing the so-called United States–China trade war. The Trump administration also changed its trade negotiation strategy from multilateralism or pluralism to bilateralism. For example, the United States withdrew from the Trans-Pacific Partnership Agreement on January 23, 2017 and suspended negotiations for the Transatlantic Trade and Investment Partnership Agreement with the EU.

Until 2016, many countries including the United States pursued deep regional economic integration by negotiating over “deep” regional trade agreements (RTAs). Lawrence (1996) distinguishes between “shallow” and “deep” integration. Shallow integration is simply trade liberalization, which involves the removal of trade barriers. By contrast, deep integration “moves beyond the removal of border barriers” (Lawrence, 1996: p. 8). Deep RTAs contain a variety of provisions including on investment, labor, the environment, and intellectual property rights (IPR). Many of the RTAs that entered into force after the mid-1990s are “deep” in this sense.

To illustrate how recent RTAs are deeper than older ones, Rodrik (2018) compares two RTAs signed by the United States: the United States–Israel free trade agreement (FTA) and the United States–Singapore FTA. Both Israel and Singapore are small nations. The United States–Israel FTA, which entered into force in 1985, was the first bilateral trade agreement that the United States signed in the postwar period. It contains only 22 articles and three annexes. Thus, it is a short agreement: less than 8,000 words in length. By contrast, the United States–Singapore FTA entered into force in 2004. It is much longer than the United States–Israel FTA. It contains 20 chapters with many articles in each and more than a dozen annexes. The total length is about 70,000 words. These two agreements are also distinct in the coverage of policy areas. Most of the articles in the United States–Israel FTA are devoted to trade liberalization issues such as the elimination of the duties and other restrictive regulations, free of import licensing requirements, and rules of origin (RoOs). On the contrary, in the United States–Singapore FTA, only the first seven chapters cover trade issues. The remaining 13 chapters address a variety of policy areas including financial services, anti-competitive business conduct, e-commerce, investment, IPR, labor, and the environment. Thus, viewing the policy areas covered by these two agreements, it is clear that the United States–Singapore FTA is deeper than the United States–Israel FTA.

Why have countries recently pursued deep regional economic integration? Lawrence (1996) argues that the development of regional production systems and promotion of service investment became important in the 1990s. To facilitate international investment and the operation of multinational enterprises, deeper forms of economic integration such as the elimination of differences in national production and product standards that make regionally integrated production costly are required.

Baldwin (2011, 2016) explains in detail why countries have been pursuing deep regional economic integration. He calls the second phase of globalization initiated by the information and communication technology revolution in the 1990s the “second unbundling,” which is the geographic separation of factories. The fragmentation of production processes and offshoring occur in the second unbundling. The information and communication technology revolution

---

1) “Regional” here does not necessarily imply geographic proximity. Geographically distant countries can form “regional” economic integrations.
enabled know-how in rich (North) countries to be combined with low-wage labor in poor (South) countries. Baldwin (2016) calls this combination the high-tech/low wage mix. In this era of the second unbundling, the so-called trade-investment-services-intellectual property nexus (Baldwin, 2016) emerged because of this mix. That is, multinational enterprises from rich countries bring their intangible property to the factories built in poor countries and conduct part of the production process there. Parts and components are traded between parent firms and their foreign affiliates. Thus, the trade of goods, the movement of capital, services that connect unbundled factories, and intellectual property are all involved in the production of branded goods under the above nexus. Baldwin (2016) argues that this nexus thus requires a new package of disciplines, which can be provided by deep RTAs.

Although a number of studies have argued that recent RTAs are deeper than older ones, no previous studies have systematically examined such a trend. Moreover, Lawrence (1996) argues that deep integration could be better or worse than shallow integration. According to him, deep integration could “take the form of imposing measures on countries that are inappropriate for their stage of development, such as excessively stringent environmental standards, or which reduce economic efficiency” (p. 8). Thus, it is important to analyze the effects of deep RTAs on members’ economies.

Taking these issues into account, in Section 2, I first examine whether recent RTAs are deeper than older ones using data on the contents of RTAs. I show that recent RTAs are deeper than older ones, whereas there is heterogeneity by signatory groups. Moreover, the provisions on import tariffs and non-tariff barriers are covered by most of the RTAs that entered into force after 2010, whereas the coverage of the provisions on the behind-the-border policies and other policies is still limited for RTAs that entered into force even after 2010. I further examine how RTAs are heterogeneous in terms of the depth and breadth according to types of RTAs, country pairs, and particular signatory countries. I show that RTAs are actually heterogeneous according to these characteristics, but differences are limited to relatively small number of categories of the depth and breadth measures. Moreover, the differences have recently narrowed. Then, I explain the effects of deep RTAs by reviewing a number of recent studies in Section 3. As the effects of deep RTAs, I focus on the trade in goods; foreign direct investment (FDI), production networks, and licensing; and technology spillovers and research collaborations. Reviewing the findings of existing studies, the effects of deep RTAs are clarified. Finally, in Section 4, I provide concluding remarks.

2 Characteristics of Deep RTAs

2.1 Are recent RTAs deeper than older ones?

In this subsection, I examine whether RTAs that were signed and entered into force in the 2000s are deeper than those before 2000. To do so, I need information on the contents of RTAs. Horn et al. (2010) propose a systematic method to measure the depth and nature of the economic integration of RTAs, which examines the policy areas covered by their provisions and legal enforceability of these obligations. They identify 52 policy areas covered by RTAs of which either the United States or the European Community is a member, classifying them into two groups: WTO-plus (WTO+) and WTO-extra (WTO-X). The WTO+ group includes provisions that fall under the current mandate of the WTO, but go beyond commitments at the
What does Regional Economic Integration Deliver?

Table 1: List of WTO+ and WTO-X Areas in RTAs

| WTO+ Areas (14)                  | WTO-X Areas (38)                        |
|---------------------------------|----------------------------------------|
| FTA Industrial Goods            | Anti-Corruption                        |
| FTA Agricultural Goods          | Competition Policy                     |
| Customs Administration          | Environmental Laws                     |
| Export Taxes                    | IPR                                     |
| SPS                             | Investment                              |
| TBT                             | Labor Market Regulation                 |
| State Trading Enterprises       | Movement of Capital                     |
| Anti-dumping                    | Consumer Protection                     |
| Countervailing Measures         | Data Protection                         |
| State Aid                       | Agriculture                             |
| Public Procurement              | Approximation of Legislation            |
| TRIMs                           | Audio Visual                            |
| GATS                            | Civil Protection                        |
| TRIPs                           | Innovation Policies                    |
| TRIPs                           | Cultural Cooperation                    |
|                       | Economic Policy Dialogue                |
|                       | Education and Training                  |
|                       | Energy                                   |
|                       | Financial Assistance                    |
|                       | Anti-dumping                             |
|                       | Competition Policy                      |
|                       | Environmental Laws                      |
|                       | IPR                                     |
|                       | Investment                              |
|                       | Labor Market Regulation                 |
|                       | Movement of Capital                     |
|                       | Agriculture                             |
|                       | Approximation of Legislation            |
|                       | Audio Visual                            |
|                       | Civil Protection                        |
|                       | Innovation Policies                    |
|                       | Cultural Cooperation                    |
|                       | Economic Policy Dialogue                |
|                       | Education and Training                  |
|                       | Energy                                   |
|                       | Financial Assistance                    |

Source: Horn et al. (2010).

multilateral level. By contrast, WTO-X provisions include issues that fall outside the current WTO mandate. The WTO+ group includes 14 provisions and WTO-X group includes 38 provisions, as shown in Table 1.

Hofmann et al. (2019) substantially extend the coverage of RTAs in the dataset presented by Horn et al. (2010) to 279.2) The extended dataset, which covers 1958–2015, is provided on the World Bank’s website.3)

The method developed by Horn et al. (2010) evaluates the coverage and legal enforceability of each provision in RTAs using two indexes: the AC and LE indexes. The AC index simply indicates whether a provision is covered by an RTA. It takes the value 1 if a provision is mentioned and 0 otherwise. On the contrary, the LE index evaluates the legal enforceability of each provision on a three-point scale: 0 for not mentioned in the agreement or not legally enforceable, 1 for legally enforceable but explicitly excluded by a dispute settlement provi-

2) An alternative database on the contents of RTAs called DESTA (Design of Trade Agreements) was created by Dür et al. (2014). It originally covered 587 trade agreements for 1945–2009. This database has now been updated and extended to more than 1,000 trade agreements. The period has also been extended to 2018. Although only 16 policy areas (10 trade and six non-trade areas) are covered by DESTA, each policy area is evaluated in detail by multiple variables.

3) URL: https://datacatalog.worldbank.org/dataset/content-deep-trade-agreements.
sion, and 2 for legally enforceable.

Figures 1 (a) and (b) show the average number of WTO+ and WTO-X provisions included in RTAs by signatory group and year of entry, respectively. The signatory groups are the United States, European countries, ASEAN members, Japan, China, Russia, and other countries. The period is divided into three: until 1999, from 2000 to 2009, and from 2010 to 2015. For example, as shown in Figure 1 (a), the RTAs signed by the United States before 2000 include, on average, 12 WTO+ provisions (of the 14 WTO+ provisions) with the LE index being either 1 or 2; those in 2000-2009 include, on average, 12.1 WTO+ provisions with \( LE \geq 1 \); and those in 2010-2015 include, on average, 12.7 WTO+ provisions with \( LE \geq 1 \). The numbers placed above the bars indicate the number of RTAs signed by the signatory country/group that entered into force in each period. For example, the United States signed two RTAs before 2000, nine RTAs during 2000-2009, and three RTAs during 2010-2015.

This figure indicates that the RTAs signed by the United States cover most of the WTO+ provisions with at least some legal enforceability even before 2000; on average, more than 12 provisions of the 14 are covered and legally enforceable. The RTAs signed by the EU included fewer than 10 WTO+ provisions, on average, until 2009 but more (11.6 on average) in the 2010s. Although the RTAs formed by ASEAN countries and Russia before 2000 included a small number of WTO+ provisions (2.4 and 4.9 on average, respectively), those in the 2000s included more WTO+ provisions (about 10 provisions on average). As for Japan and China, the number of legally enforceable WTO+ provisions in RTAs increased in the 2010s: 12 (Japan) and 11.6 (China) provisions on average.\(^4\) The number of legally enforceable WTO+ provisions included in the RTAs signed by other countries has also steadily increased.

By contrast, panel (b) of Figure 1 shows the average number of WTO-X provisions included in RTAs with the LE index taking the value of \( LE \geq 1 \). Although there are 38

---

\( ^4 \) Both Japan and China signed no RTAs before 2000.

---
WTO-X provisions, the average number included in RTAs with at least some legal enforceability is fewer than nine for all countries, even after 2010. In general, recent RTAs include more legally enforceable WTO-X provisions for all countries; however, the highest average is still around eight provisions in the 2010s (8.3 for the United States, 8.4 for the EU, and 8.0 for Japan).

In summary, Figure 1 suggests that recent RTAs cover most of the provisions that fall under the current mandate of the WTO with legal enforceability, but still address only a limited number of the policy areas outside the WTO discipline.

Limão (2016) proposes recategorizing the WTO+ and WTO-X provisions from the viewpoint of the depth and breadth of RTAs (see Table 2). First, the depth of RTAs is measured by four categories of provisions in the WTO+ and WTO-X groups: (a) import tariffs, (b) non-tariff barriers (NTBs), (c) behind-the-border policies (BBPs), and (d) other policies. Category (a) includes two provisions, category (b) includes six provisions, category (c) consists of five provisions, and category (d) includes 16 provisions. Second, the breadth of RTAs is measured by five categories: (a) services, (b) technology, (c) investment/capital, (d) labor, and (e) non-economic policies (non-EPs). Category (a) includes one provision, category (b) consists of six provisions, category (c) has three provisions, category (d) four provisions, and category (e) nine provisions.

Figure 2 shows the depth of RTAs by signatory group and year of entry. Each bar indicates the shares of $LE=1$ (light color) and $LE=2$ (dark color) provisions included in each category in all the RTAs signed by each country/group that entered into force in each period. As shown in panel (a), most RTAs, particularly those that entered into force in the 2000s, fully cover both provisions in the category of import tariffs with $LE=2$. The coverage of provisions in the non-tariff barriers category is also relatively high. More than 80% of the RTAs signed by the United States before 2000 cover the provisions of non-tariff barriers with $LE=2$. Since 2010, about 90% of the RTAs signed by the United States cover those provisions, but about 40% are with $LE=1$. The RTAs signed by the EU, Japan, and Russia in 2010–2015 also cover about 90% of the provisions in this category. The coverage of the RTAs signed by China in the 2000s is much higher and close to 1. The coverage of the provisions in the behind-the-border policies category is rather low. Even the highest level of coverage is about 0.7 for the RTAs signed by the United States and the EU in 2010–2015. The coverage of the provisions in this category by the RTAs signed by ASEAN countries, China, and Russia is less than 0.5. Moreover, the coverage of the provisions in the category of other policies is much lower. The coverage is below 0.2 for all RTAs in all periods.

Next, Figure 3 shows the breadth of RTAs by signatory country/group and year of entry. As in Figure 2, each bar indicates the shares of $LE=1$ (light color) and $LE=2$ (dark color) provisions included in each category in all the RTAs signed by each country/group that entered into force in each period.

The categories for which the coverage of provisions is relatively high are (a) services and (c) investment/capital. However, coverage depends on the signatory country. The RTAs signed by the United States, Japan, and ASEAN countries in the 2000s cover categories (a) and (c) at relatively high shares. The coverage of the service area in the RTAs signed by China after 2010 is 100%. By contrast, the RTAs signed by the EU and Russia only cover provisions in categories (a) and (c) at low shares. The coverage of other categories ((b), (d), and (e))
tends to be low for all RTAs. In particular, the coverage of (e) (non-economic policies) is less than 0.2 for all RTAs.

### 2.2 Heterogeneity in the depth and breadth of RTAs

In the literature, a number of studies have found the heterogeneous effects of RTAs by the type of agreements (Roy, 2010; Vicard, 2009) and by the characteristics of country pairs and

#### Table 2: Depth and Breadth of RTAs

| Depth Category | Provisions |
|----------------|------------|
| (a) Import tariffs | FTA Industrial Goods |
| | FTA Agricultural Goods |
| (b) Non-tariff barriers | Customs Administration |
| | Export Taxes |
| | SPS |
| | TBT |
| | Anti-dumping |
| | Countervailing Measures |
| (c) Behind the border policies | State Trading Enterprises |
| | State Aid |
| | Public Procurement |
| | Anti-Corruption |
| | Competition Policy |
| (d) Other policies | Consumer Protection |
| | Data Protection |
| | Agriculture |
| | Approximation of Legislation |
| | Civil Protection |
| | Education and Training |
| | Energy |
| | Financial Assistance |
| | Industrial Cooperation |
| | Mining |
| | Nuclear Safety |
| | Public Administration |
| | Regional Cooperation |
| | SME |
| | Statistics |
| | Taxation |
| (e) Non-economic policies | Environmental Laws |
| | Audio Visual |
| | Cultural Cooperation |
| | Health |
| | Human Rights |
| | Illicit Drugs |
| | Money Laundering |
| | Political Dialogue |
| | Terrorism |

| Breadth Category | Provisions |
|------------------|------------|
| (a) Services | GATS |
| (b) Technology | TRIPs |
| | IPR |
| | Innovation Policies |
| | Economic Policy Dialogue |
| | Information Society |
| | Research and Technology |
| (c) Investment /capital | TRIMs |
| | Investment |
| | Movement of Capital |
| (d) Labor | Labor Market Regulation |
| | Illegal Immigration |
| | Social Matters |
| | Visa and Asylum |
| (e) Non-economic policies | Environmental Laws |
| | Audio Visual |
| | Cultural Cooperation |
| | Health |
| | Human Rights |
| | Illicit Drugs |
| | Money Laundering |
| | Political Dialogue |
| | Terrorism |

*Source: Limão (2016).*
What does Regional Economic Integration Deliver?

Based on these findings by the existing studies, I examine how the heterogeneity in the depth and breadth of RTAs is affected by the characteristics of RTAs. As the sources of the heterogeneity, I focus on (i) the type of RTAs (i.e., customs union and free trade agreement), (ii) particular signatory countries, and (iii) the type of country pairs (i.e., North-North, North-South, and South-South agreements).

I first analyze the heterogeneity by the type of RTAs. I divide the type of RTAs into three: customs union (CU), free trade agreement (FTA), and partial scope agreement (PSA).\(^6\) In the analysis below, I mainly focus on the differences between CUs and FTAs. To compare how CUs and FTAs are different in the depth and breadth, I calculate the mean and the stan-

\(^5\) I thank an anonymous referee for suggesting this important issue.

\(^6\) Both CUs and FTAs may be combined with economic integration agreements (EIAs).
standard deviation of the number of \( LE=1 \) or \( LE=2 \) provisions for each category of the depth and breadth indexes. Moreover, to examine how recent RTAs are different from older one, I calculate the mean and the standard deviation for both all period (1957-2015) and the recent period (2000-2015). The results are reported in Tables 3 (the depth categories) and 4 (the breadth categories).

The statistical significance of the differences in the mean among the types of RTAs is tested by the one-way analysis of variance (ANOVA). The value of \( F \) statistic and the corresponding \( p \)-value are reported at the bottom of the table.

As for the depth indexes, the heterogeneity in import tariffs is statistically significant at 5% in all period, whereas it is insignificant in the recent period. The heterogeneity in the other categories of the depth measure is statistically significant in both all period and the recent period. However, if we focus on the difference between CUs and FTAs, the direct comparison between these two indicates that CUs are deeper than FTAs in the category of other policies in both all period and the recent period (\( p<0.001 \)), whereas the average value of the import tariffs provisions for FTAs is higher than that for CUs in all period (\( p<0.028 \)).\(^7\) The difference between CUs and FTAs is not statistically significant for both the NTBs and BBPs provisions.

With regard to the breadth indexes, although the heterogeneity among three types of

\(^7\) For brevity, the results of the statistical tests for the direct comparison between two groups are not reported in tables. The same applies to the other tables.
What does Regional Economic Integration Deliver?

Table 3: Heterogeneity in the Depth Categories by RTA Types

|                  | All period (1957–2015) | Recent period (2000–2015) |
|------------------|------------------------|--------------------------|
|                  | No. of RTAs | Tariffs | NTBs | BBPs | Other | No. of RTAs | Tariffs | NTBs | BBPs | Other |
| CU               | 29         | 1.862   | 4.103 | 2.034 | 3.276 | 13          | 2.000   | 4.846 | 1.846 | 3.077 |
|                  |            | (0.516) | (2.177) | (1.658) | (4.026) |            | (0.000) | (1.908) | (1.725) | (4.132) |
| FTA              | 233        | 1.987   | 4.069 | 2.356 | 0.725 | 178         | 1.989   | 4.584 | 2.517 | 0.798 |
|                  |            | (0.146) | (1.741) | (1.292) | (1.554) |            | (0.145) | (1.487) | (1.281) | (1.741) |
| PSA              | 17         | 1.882   | 1.294 | 0.529 | 0.647 | 7           | 2.000   | 1.714 | 0.714 | 0.000 |
|                  |            | (0.485) | (1.897) | (1.068) | (2.242) |            | (0.000) | (2.059) | (0.756) | (0.000) |
| Total            | 279        | 1.968   | 3.903 | 2.211 | 0.986 | 198         | 1.990   | 4.500 | 2.409 | 0.919 |
|                  |            | (0.245) | (1.912) | (1.389) | (2.144) |            | (0.142) | (1.621) | (1.344) | (2.029) |

ANOVA

\[ \begin{align*}
F (2, 276) &= 4.57 \quad (p < 0.012) \quad (p < 0.001) \quad (p < 0.001) \quad (p < 0.001) \\
F (2, 195) &= 0.06 \quad 12.10 \quad 7.78 \quad 9.08 \quad (p < 0.9456) \quad (p < 0.001) \quad (p < 0.001) \quad (p < 0.001)
\end{align*} \]

Notes: (a) The mean of the number of \( LE = 1 \) or \( LE = 2 \) provisions for each category of the depth indexes is calculated by type of RTAs. The number in parentheses is the corresponding standard deviation. (b) The number of provisions included in each category (which is equal to the highest number of the index) is 2 (tariffs), 6 (NTBs), 5 (BBPs), and 16 (other), respectively. (c) The value of \( F \) statistic and the corresponding \( p \)-value to test the differences in the mean among all groups by ANOVA are reported at the bottom.

RTAs is statistically significant for all categories, the difference between CUs and FTAs is significant only for labor provisions and non-EP provisions. In both categories, the average value of the index is higher for CUs than for FTAs.

I next examine the heterogeneity by participation of particular countries. As the groups of countries, I focus on the United States, European countries, Japan, China, and Russia.8) The results are reported in Tables 5 (the depth categories) and 6 (the breadth categories).

In terms of the depth categories, there is no heterogeneity among five country groups for import tariffs provisions and other policies provisions. Moreover, the heterogeneity is not statistically significant for NTBs provisions in 2000–2015. However, by comparing two country groups, I find that the average value of the NTBs index for RTAs signed by Russia is significantly lower than that for RTAs signed by the Unites States and European countries \((p < 0.01)\) and also lower than that for RTAs signed by Japan \((p < 0.10)\) in all period. RTAs signed by Russia are also lower than those signed by the United States and Europe for the BBPs index in all period \((p < 0.001)\) and in the recent period \((p < 0.02)\). Moreover, RTAs signed by China are lower than those signed by the United States and Europe for the BBPs index \((p < 0.001)\) in both all period and 2000–2015.

In terms of the breadth categories, on the other hand, the heterogeneity is statistically significant for the investment/capital provisions. In particular, the average value of the investment/capital indexes for RTAs signed by China and Russia are significantly lower than those

8) Unlike the analysis in the previous subsection, I exclude the groups of ASEAN countries and other countries. This is because each group must be mutually exclusive.
Table 4: Heterogeneity in the Breadth Categories by RTA Types

|                  | All period (1957-2015) | Recent period (2000-2015) |
|------------------|-------------------------|--------------------------|
|                  | No. of RTAs | Services | Tech | Inv/Cap | Labor | NonEPs | No. of RTAs | Services | Tech | Inv/Cap | Labor | NonEPs |
| CU               | 29          | 0.483    | 1.448| 1.207    | 1.310  | 1.241   | 13          | 0.385    | 1.615| 1.231    | 1.231  | 1.538  |
|                  |             | (0.509)  | (1.502)| (0.940)  | (1.692)| (1.480) |             | (0.506)  | (1.710| (0.927)  | (1.481)| (1.391)|
| FTA              | 233         | 0.536    | 1.206| 1.270    | 0.506  | 0.361   | 178         | 0.652    | 1.461| 1.511    | 0.612  | 0.455  |
|                  |             | (0.500)  | (1.099)| (1.211)| (0.677)| (0.960) |             | (0.478)  | (1.063)| (1.194)  | (0.698)| (1.074)|
| PSA              | 17          | 0.118    | 0.235| 0.176    | 0.118  | 0.294   | 7           | 0.000    | 0.000| 0.000    | 0.000  | 0.000  |
|                  |             | (0.332)  | (0.970)| (0.529)| (0.485)| (1.213) |             | (0.000)  | (0.000)| (0.000)  | (0.000)| (0.000)|
| Total            | 279         | 0.505    | 1.172| 1.197    | 0.566  | 0.448   | 198         | 0.611    | 1.419| 1.439    | 0.631  | 0.510  |
|                  |             | (0.501)  | (1.163)| (1.182)| (0.870)| (1.071) |             | (0.489)  | (1.127)| (1.190)  | (0.781)| (1.111)|
| ANOVA            | F(2, 276)  | 5.76     | 6.70 | 7.09     | 14.74  | 9.45    | F(2, 195)   | 8.01     | 6.18 | 5.93     | 6.52   | 6.91   |
|                  |             | (p < 0.004) | (p < 0.002) | (p < 0.001) | (p < 0.001) | (p < 0.001) |             | (p < 0.001) | (p < 0.003) | (p < 0.004) | (p < 0.002) | (p < 0.002) |

Notes: (a) The mean of the number of LE = 1 or LE = 2 provisions for each category of the breadth indexes is calculated by type of RTAs. The number in parentheses is the corresponding standard deviation. (b) The number of provisions included in each category (which is equal to the highest number of the index) is 1 (services), 6 (technology), 3 (investment/capital), 4 (labor), and 9 (non-EPs), respectively. (c) The value of F statistic and the corresponding p-value to test the differences in the mean among all groups by ANOVA are reported at the bottom.
What does Regional Economic Integration Deliver?

Table 5: Heterogeneity in the Depth Categories by Particular Countries

| Country     | All period (1957–2015) | Recent period (2000–2015) |
|-------------|-------------------------|---------------------------|
|             | No. of RTAs | Tariffs | NTBs | BBPs | Other | No. of RTAs | Tariffs | NTBs | BBPs | Other |
| US          | 14          | 2.000   | 5.214 | 3.214 | 0.357 | 12          | 2.000   | 5.250 | 3.250 | 0.333 |
|             | (0.000)     | (1.477) | (1.251) | (0.633) | | (0.000)     | (1.545) | (1.288) | (0.651) | |
| Europe      | 72          | 2.000   | 4.846 | 3.056 | 1.792 | 50          | 2.000   | 4.660 | 3.140 | 1.420 |
|             | (0.000)     | (1.321) | (1.221) | (3.642) | | (0.000)     | (1.206) | (1.195) | (3.381) | |
| Japan       | 14          | 2.000   | 4.500 | 2.500 | 1.500 | 14          | 2.000   | 4.500 | 2.500 | 1.500 |
|             | (0.000)     | (1.345) | (1.160) | (2.103) | | (0.000)     | (1.344) | (1.160) | (2.103) | |
| China       | 12          | 2.000   | 4.250 | 1.083 | 0.917 | 12          | 2.000   | 4.250 | 1.083 | 0.917 |
|             | (0.000)     | (2.050) | (1.084) | (1.311) | | (0.000)     | (2.050) | (1.084) | (1.311) | |
| Russia      | 20          | 1.900   | 3.100 | 1.450 | 0.700 | 6           | 2.000   | 5.500 | 1.333 | 1.177 |
|             | (0.447)     | (1.944) | (0.887) | (1.418) | | (0.000)     | (0.837) | (1.211) | (2.041) | |
| Total       | 132         | 1.985   | 4.333 | 2.591 | 1.364 | 94          | 2.000   | 4.713 | 2.681 | 1.213 |
|             | (0.174)     | (1.605) | (1.370) | (2.898) | | (0.000)     | (1.396) | (1.401) | (2.696) | |

ANOVA

\[
F(4, 127) = 1.42, F(4, 89) = 1.37, F(4, 89) = 9.96, F(4, 89) = 0.46
\]

Notes: (a) The mean of the number of LE = 1 or LE = 2 provisions for each category of the depth indexes is calculated by signatory country of RTAs. The number in parentheses is the corresponding standard deviation. (b) The number of provisions included in each category (which is equal to the highest number of the index) is 2 (tariffs), 6 (NTBs), 5 (BBPs), and 16 (other), respectively. (c) The value of \( F \) statistic and the corresponding \( p \)-value to test the differences in the mean among all groups by ANOVA are reported at the bottom.

Finally, I investigate the heterogeneity by the type of country pairs. The results are reported in Tables 7 (the depth categories) and 8 (the breadth categories).

With regard to the depth categories, the average values of the NTBs, BBPs, and other policies indexes for the North–North agreements are significantly higher than those for the other country pairs, except for the comparison with the North–South agreements for the BBPs index in 2000–2015. Although the difference between the North–South and South–South agreements is statistically significant for both NTBs and BBPs indexes in all period, the difference becomes insignificant for the NTBs index in 2000–2015.

As for the breadth categories, the heterogeneity among the three types of country pairs is

9) The \( p \)-value is as follows: \( p<0.003 \) for all comparisons, except for the comparison between the United States and Russia, \( p<0.011 \).

10) The \( p \)-value is \( p<0.01 \) for the comparison with the North–South agreements in all period and for the comparison with the South–South agreements in both all period and in 2000–2015. As for the comparison with the North–South agreements in 2000–2015, \( p<0.02 \).
**Table 6: Heterogeneity in the Breadth Categories by Particular Countries**

|                  | All period (1957–2015) |                |                |                |                |                | Recent period (2000–2015) |                |                |                |                |                |
|------------------|------------------------|----------------|----------------|----------------|----------------|----------------|------------------------|----------------|----------------|----------------|----------------|----------------|
|                  | No. of RTAs | Services | Tech | Inv/Cap | Labor | NonEPs | No. of RTAs | Services | Tech | Inv/Cap | Labor | NonEPs |
| US               | 14          | 0.929    | (0.267) | 1.929 | (1.158) | 2.429 | (0.392) | 1.071 | (0.475) | 12          | 1.000 | (0.000) | 2.000 | (0.000) | 2.500 | (1.168) | 1.083 | (0.289) | 1.167 | (0.389) |
| Europe           | 72          | 0.444    | (0.500) | 1.806 | (1.158) | 1.125 | (0.903) | 0.861 | (1.314) | 50          | 0.460 | (0.503) | 1.900 | (1.093) | 1.120 | (0.849) | 0.700 | (1.093) | 0.700 | (1.764) |
| Japan            | 14          | 1.000    | (0.000) | 2.071 | (0.730) | 2.643 | (0.745) | 0.643 | (0.497) | 14          | 1.000 | (0.000) | 2.071 | (0.730) | 2.643 | (0.745) | 0.643 | (0.497) | 0.286 | (0.611) |
| China            | 12          | 0.667    | (0.492) | 1.583 | (1.564) | 1.083 | (0.793) | 0.667 | (0.778) | 12          | 0.667 | (0.492) | 1.583 | (1.564) | 1.083 | (0.793) | 0.667 | (0.778) | 0.417 | (0.669) |
| Russia           | 20          | 0.200    | (0.410) | 0.450 | (0.759) | 0.400 | (0.754) | 0.250 | (0.550) | 6           | 0.500 | (0.548) | 1.333 | (1.095) | 1.000 | (1.095) | 0.500 | (1.548) | 0.833 | (1.329) |
| Total            | 132         | 0.538    | (0.500) | 1.621 | (1.156) | 1.311 | (1.113) | 0.742 | (1.060) | 94          | 0.638 | (0.483) | 1.861 | (1.033) | 1.511 | (1.095) | 0.723 | (0.885) | 0.670 | (1.386) |

**ANOVA**

\[
F(4, 127) = 10.64 \quad (p < 0.001) \\
F(4, 89) = 6.86 \quad (p < 0.001)
\]

Notes: (a) The mean of the number of LE = 1 or LE = 2 provisions for each category of the breadth indexes is calculated by signatory country of RTAs. The number in parentheses is the corresponding standard deviation. (b) The number of provisions included in each category (which is equal to the highest number of the index) is 1 (services), 6 (technology), 3 (investment/capital), 4 (labor), and 9 (non-EPs), respectively. (c) The value of F statistic and the corresponding p-value to test the differences in the mean among all groups by ANOVA are reported at the bottom.
What does Regional Economic Integration Deliver?

Table 7: Heterogeneity in the Depth Categories by Pair Types

|                  | All period (1957–2015) | Recent period (2000–2015) |
|------------------|-------------------------|---------------------------|
|                  | No. of RTAs | Tariffs | NTBs | BBPs | Other | No. of RTAs | Tariffs | NTBs | BBPs | Other |
| North-North      | 31          | 2.000   | 5.387 | 3.419 | 3.161 | 17          | 2.000   | 5.706 | 3.412 | 2.353 |
|                  | (0.000)     | (1.022) | (0.923)| (4.148)|        | (0.000)     | (0.588)| (1.004)| (3.707)|        |
| North-South      | 132         | 2.000   | 4.288 | 2.689 | 0.811 | 112         | 2.000   | 4.563 | 2.795 | 0.893 |
|                  | (0.000)     | (1.470) | (1.285)| (1.891)|        | (0.000)     | (1.341)| (1.224)| (2.020)|        |
| South-South      | 116         | 1.922   | 3.069 | 1.345 | 0.603 | 69          | 1.971   | 4.101 | 1.536 | 0.609 |
|                  | (0.377)     | (2.154) | (1.072)| (1.046)|        | (0.241)     | (2.016)| (1.132)| (1.191)|        |
| Total            | 279         | 1.968   | 3.903 | 2.211 | 0.986 | 198         | 1.990   | 4.500 | 2.409 | 0.919 |
|                  | (0.245)     | (1.912) | (1.389)| (2.144)|        | (0.142)     | (1.621)| (1.344)| (2.029)|        |

ANOVA

F (2, 276) 3.45  | 27.43  | 59.96  | 20.85  |
(p < 0.033) | (p < 0.001) | (p < 0.001) | (p < 0.001) |

F (2, 195) 0.93  | 7.32   | 31.22  | 5.28   |
(p < 0.395) | (p < 0.001) | (p < 0.001) | (p < 0.006) |

Notes: (a) The mean of the number of LE = 1 or LE = 2 provisions for each category of the depth indexes is calculated by type of country pairs. The number in parentheses is the corresponding standard deviation. (b) The number of provisions included in each category (which is equal to the highest number of the index) is 2 (tariffs), 6 (NTBs), 5 (BBPs), and 16 (other), respectively. (c) The value of F statistic and the corresponding p-value to test the differences in the mean among all groups by ANOVA are reported at the bottom.

particularly evident for the technology index: The mean value of the technology index is the highest for the North–North agreements, the second highest for the North–South agreements, and the lowest for the South–South agreements in both all period and 2000–2015.11) Although the difference in the investment/capital and labor indexes between the North–North agreements and the South–South agreements is significant in all period (p < 0.001), the differences for the other comparisons are insignificant or significant but the p-value is higher.

In sum, CUs tend to be deeper than FTAs in the index of other policies in the depth category and in the index of non-EPs in the breadth category. On the other hand, RTAs signed by Russia are on average shallower than those signed by the United States and European countries in the indexes of NTBs and BBPs in the depth category and shallower than those signed by the United States and Japan in the indexes of services, technology, and investment/capital in the breadth category. Similarly, RTAs signed by China are on average shallower than those signed by the United States and European countries in the index of BBPs in the depth category and shallower than those signed by the United States and Japan in the index of investment/capital in the breadth category. Finally, although North–North RTAs tend to be deeper than RTAs by other types of country pairs in the indexes of NTBs, BBPs, and other policies in the depth category, the differences have narrowed in the 2000s. Moreover, in terms of the breadth of provisions, the ordering of the North–North, North–South, and South–South RTAs is significant for the index of technology, the heterogeneity among those three types of RTAs is less evident for other aspects of the breadth, particularly in the 2000s.

11) The p-value for the comparison between the North–North and North–South agreements in 2000–2015 is p < 0.012 and for the other comparisons p < 0.001.
### Table 8: Heterogeneity in the Breadth Categories by Pair Types

|                     | All period (1957-2015) | Recent period (2000-2015) |
|---------------------|-------------------------|---------------------------|
|                     | No. of RTAs | Services | Tech | Inv/Cap | Labor | NonEPs | No. of RTAs | Services | Tech | Inv/Cap | Labor | NonEPs |
| North-North         | 31          | 0.742     | 2.258| 1.774 | 1.548 | 1.452  | 17          | 0.824     | 2.411| 2.059 | 1.235 | 1.353 |
|                     | (0.445)     | (1.125)   | (1.175) | (1.567) | (1.588) |        | (0.393)     | (1.004)   | (1.144) | (1.348) | (1.412) |
| North-South         | 132         | 0.553     | 1.492| 1.364 | 0.568 | 0.439  | 112         | 0.598     | 1.643| 1.438 | 0.634 | 0.500 |
|                     | (0.499)     | (1.081)   | (1.161) | (0.723) | (1.114) |        | (0.492)     | (1.047)   | (1.169) | (0.747) | (1.193) |
| South-South         | 116         | 0.388     | 0.517| 0.853 | 0.302 | 0.190  | 69         | 0.580     | 0.812| 1.290 | 0.478 | 0.319 |
|                     | (0.489)     | (0.860)   | (1.113) | (0.514) | (0.603) |        | (0.497)     | (0.974)   | (1.202) | (0.559) | (0.757) |
| Total               | 279         | 0.505     | 1.172| 1.197 | 0.566 | 0.448  | 198        | 0.611     | 1.419| 1.439 | 0.631 | 0.510 |
|                     | (0.501)     | (1.163)   | (1.182) | (0.870) | (1.071) |        | (0.489)     | (1.127)   | (1.190) | (0.781) | (1.112) |

**ANOVA**

\[
F(2, 276) = 7.59 \quad \text{(p < 0.001)}
\]

\[
F(2, 195) = 1.80 \quad \text{(p < 0.168)}
\]

*Notes:* (a) The mean of the number of LE = 1 or LE = 2 provisions for each category of the breadth indexes is calculated by type of country pairs. The number in parentheses is the corresponding standard deviation. (b) The number of provisions included in each category (which is equal to the highest number of the index) is 1 (services), 6 (technology), 3 (investment/capital), 4 (labor), and 9 (non-EPs), respectively. (c) The value of $F$ statistic and the corresponding p-value to test the differences in the mean among all groups by ANOVA are reported at the bottom.
3 The effects of deep RTAs

As cited in Section 1, Lawrence (1996) argues that deep RTAs could be better or worse than shallow ones. Thus, it is important to examine the effects of deep RTAs. A number of studies have investigated the effects of deep RTAs from various aspects. In this section, I briefly review studies of the effects of deep RTAs on (i) trade in goods; (ii) FDI, production networks, and licensing; and (iii) technology spillovers and research collaborations.

3.1 The effects of deep RTAs on trade in goods

Examining the effects of deep RTAs on trade in goods is a natural extension of the traditional study of the economic effects of RTAs by Viner (1950). That is, the focus is on the trade creation and trade diversion effects. Mattoo et al. (2017) investigate how deep integration alters the trade creation and trade diversion effects of RTAs. They explain differences in the trade creation and trade diversion effects between shallow and deep RTAs using a simple model of three symmetric countries (Home, Partner, and RoW) with three goods, in which each country exports two goods and imports one good. In addition to a specific tariff imposed by an importing country on each good, there are frictional trade barriers created by non-tariff measures. Then, Home and Partner sign an RTA, either shallow or deep one. While a shallow RTA would only eliminate the tariff between members, a deep RTA eliminates both the tariff and the frictional barriers between them, which results in making the trade creation larger. Moreover, unlike the shallow RTA, the deep RTA could reduce part of the frictional barriers that non-members face in exporting goods to members. Although the effect is ambiguous, the trade diversion effect could be smaller. Now, based on this theoretical foundation, Mattoo et al. (2017) empirically examine how deep integration alters the trade creation and trade diversion effects of RTAs. They measure the depth of RTAs using three alternative indexes: (a) the count of legally enforceable provisions included in each RTA, (b) the count of the areas covered irrespective of their legal enforceability, and (c) the count of provisions more likely to be economically relevant. All these indexes are normalized between 0 and 1. The estimated coefficient of the depth index captures the trade creation effect of deep RTAs. On the contrary, the trade diversion effect is estimated as the coefficient of the variable that measures the importer’s average depth of RTAs with the rest of the world (i.e., all trading partners excluding the exporting country). They find that deep RTAs lead to greater trade creation and smaller trade diversion than shallow RTAs.

Jinji and Kamata (2020) focus on a specific aspect of deep RTAs, namely labor clauses, and examine how the labor clauses in deep RTAs affect the trade in manufacturing goods. In particular, they examine whether labor clauses reduce the trade creation effect of RTAs. They argue that labor clauses in deep RTAs may increase or decrease the trade creation effect of RTAs. They pay attention to the negotiation stage of RTAs. In principle, when countries negotiate over RTAs, labor clauses impose on the negotiating parties a policy burden by urging them to commit to adjusting their domestic policies. The burden may be larger when a country needs to harmonize its domestic labor conditions with the partner that has higher labor standards. In such a case, the country may be conservative in giving the partner access to its market and try not to lower tariffs and other trade barriers much against the partner, resulting in a
smaller trade creation effect. However, there is another possibility in which the country may give more generous market access to its RTA partner in return for the partner’s commitment to the extra policy burden by applying higher labor standards. In this case, the trade creation effect could be larger compared with RTAs without labor clauses. Then, Jinji and Kamata (2020) employ a standard structural gravity framework and find that legally enforceable labor clauses tend to increase the trade creation effect, whereas legally unenforceable labor clauses may reduce it. They also show that these impacts on the trade creation effect are heterogeneous among types of country pairs. For example, the positive effect of legally enforceable labor clauses is statistically significant for trade between Northern countries but insignificant for trade between Southern countries. Moreover, they find that once the endogeneity of labor clauses is addressed by the instrumental variable estimation, all the estimated impacts become insignificant. This result suggests that countries may selectively choose the inclusion of labor clauses in the negotiation of RTAs by anticipating their impact on trade with partner countries.

3.2 The effects of deep RTAs on FDI, production networks, and licensing

Several studies investigate the impact of deep RTAs on FDI, production networks (or global value chains; GVCs), and licensing. As argued in Section 1, the main reason for countries to pursue deep RTAs is that the trade-investment-services-intellectual property nexus for GVCs requires a new package of disciplines provided by deep RTAs (Baldwin, 2016). FDI and licensing are also involved in the nexus. Therefore, interest in the impact of deep RTAs on FDI, GVCs, and licensing naturally emerges.

First, Osnago et al. (2016, 2019) examine whether deep RTAs increase vertical FDI. Osnago et al. (2019) provide the following theoretical background about the impact of deep RTAs on vertical FDI. They employ a simplified version of the model developed by Antràs and Helpman (2004, 2008). In their model, heterogeneous firms produce a final good from two inputs: components and headquarter services. Each input is produced with a continuum of partially contractible activities. There are two countries: North and South. Contractibility is complete in the North, whereas it is incomplete in the South. Differences in contractibility across production processes and across countries reflect technological and institutional variation. The institutional environment is determined by the characteristics of domestic institutions, but it is also affected by the contents of deep RTAs (e.g., TBT, IPR, and investment provisions) that the country signs. Each final good producer decides whether to source components in the North or in the South and whether to engage in vertical FDI or foreign outsourcing. In this model setting, Osnago et al. (2019) show that deep RTAs with provisions improving the contractibility of components, such as TBT provision, increase the share of firms engaging in vertical FDI. By contrast, deep RTAs with provisions improving the contractibility of headquarter services, such as IPR and investment provisions, increase the share of firms engaging in foreign outsourcing and reduce the share of firms engaging in vertical FDI. Now, based on this theoretical background, Osnago et al. (2016, 2019) empirically analyze the effects of deep RTAs on vertical FDI. Both studies measure vertical FDI from country $i$ to country $j$ in sector $k$ using the aggregate revenue of all the subsidiaries owned by the firms in country $i$ producing inputs for sector $k$ in country $j$, using firm-level data. Osnago et al. (2016) find that even after addressing the endogeneity of deep RTAs using the instrumental variable approach, deep RTAs have significantly positive effects on vertical FDI. Osnago et al.
What does Regional Economic Integration Deliver?

(2019) extend the analysis and investigate how the contents of deep RTAs are related to vertical FDI. They find that provisions that improve the contractibility of the components provided by suppliers, such as standards and other regulatory requirements promoting harmonization and mutual recognition, are positively correlated with vertical FDI. By contrast, provisions that improve the contractibility of headquarters services such as IPR and investment protection tend to have a negative relationship with FDI. Both findings support their theoretical predictions.

Second, Orefice and Rocha (2014) analyze the relationship between deep RTAs and production networks. They first estimate an augmented gravity model with import values in parts and components being the dependent variable to capture production networks trade. They find that signing deep RTAs increases production networks trade between member countries by about 12 percentage points. They also analyze whether higher levels of production networks trade increase the likelihood of forming deep RTAs. After taking other RTA determinants into account, a 10 percentage point increase in the share of production networks trade over total trade increases the depth of an agreement by about 6 percentage points. Moreover, the probability of signing deep RTAs is higher for country pairs involved in North-South production networks and Asian countries.

Laget et al. (2020) extend the analysis and investigate how the depth of RTAs affects GVC participation. They estimate a structural gravity model and find that deep RTAs are associated with increases in the domestic and foreign value-added content of exports. On average, each additional policy area increases the domestic value added of intermediate goods and services exports by 0.48% through forward GVC linkages. On the contrary, an additional provision increases the foreign value added of intermediate goods and services exports by 0.38%, on average, through backward GVC linkages. They also find that deep RTAs improve forward linkages into more complex GVCs in the sense that exported intermediate goods cross borders twice or more.

Boffa et al. (2018) compare the extent to which deep RTAs and bilateral investment treaties stimulate and shape GVCs. They analyze the impact on various trade in value added indicators, or GVC trade, by estimating an augmented gravity model. They find that both deep RTAs and bilateral investment treaties tend to increase GVC trade. However, their transmission channels differ. Backward linkages are stimulated through both deep RTAs and bilateral investment treaties. By contrast, only deep RTAs affect forward linkages. These findings suggest that negotiating a deep RTA with investment provisions has a larger impact on value-added trade than signing a shallow RTA and a separate bilateral investment treaty.

Finally, Jinji et al. (2021) examine the impact of shallow and deep RTAs on cross-border licensing. They first derive a micro-founded gravity equation for cross-border licensing based on the model in which heterogeneous firms choose their supply modes from licensing, exports, and FDI. Their model is an extension of the model developed by Helpman et al. (2004). Jinji et al. (2021) add licensing as a supply mode to the foreign market and show that lower productivity firms engage in cross-border licensing, whereas higher productivity firms serve the foreign market either by exports or by FDI. They perform a number of comparative statics analyses, showing that deep RTAs in general and deep RTAs with provisions strengthening IPR protection increases cross-border licensing. By contrast, shallow RTAs may not enhance cross-border licensing. Then, they estimate the structural gravity model using data on bilateral
flows of licensing fees and royalties and find that deep RTAs actually enhance cross-border licensing, while the effect of shallow RTAs may be negative.

3.3 The effects of deep RTAs on technology spillovers and research collaborations

Deep RTAs may also enhance spillovers of knowledge and facilitate cross-border research collaborations among member countries of the RTAs. First, Jinji et al. (2019) examine whether deep RTAs enhance international technology spillovers using a panel of patent citation data. They argue that both shallow and deep RTAs facilitate international technology spillovers, because trade in goods is a major channel of technology spillovers and a tariff reduction through RTAs increases trade among member countries. However, this is rather an indirect linkage between RTAs and technology spillovers. A more direct link between deep RTAs and technology spillovers arises from the inclusion of provisions that stimulate technology spillovers. Some RTAs actually include provisions to encourage collaborative research projects and transfer of technology between firms in member countries. Then, from an empirical analysis, Jinji et al. (2019) find that deep RTAs enhance technology spillovers. Moreover, they analyze which provisions increase technology spillovers more. Interestingly, they find that a deep integration in a broad sense has a greater impact on technology spillovers than provisions that are more directly related to technology.

Second, Jinji et al. (2018) investigate whether deep RTAs facilitate cross-border research collaborations. They first analyze theoretically whether signing a trade agreement with deep economic integration increases firms’ incentive to engage in research collaborations using a simple duopoly model. For simplicity, they assume that there is one firm in each country. Spillovers between firms give them an incentive to engage in research collaboration. They analyze the effects of shallow and deep RTAs on research collaborations as a reduction in tariffs and an increase in the degree of spillovers, respectively. They show that a formation of a shallow RTA increases firms’ incentive to engage in research collaboration, whereas the effect of deep RTA depends on the degree of spillovers and on the degree of asymmetry in tariffs. Thus, it is an empirical question whether deep RTAs actually enhance international research collaborations. Then, Jinji et al. (2018) empirically analyze the relationship between international research collaborations and various indexes of shallow and deep RTAs. They use data on patents with multiple inventors from different countries at the United States Patent and Trademark Office as the measure of international research collaborations. Their main finding is that deeper integration is related to more active cross-border research collaborations.

4 Concluding remarks

Globalization and deep regional economic integration may slow because of the anti-globalization sentiment in many countries and protectionism policy implemented by the United States and other countries. The COVID-19 pandemic may also change the momentum of globalization and deep integration. Therefore, we should carefully watch whether the trend of deeper integration until 2016 continues.

---

12) The use of patent citation data as direct evidence of technology spillovers was pioneered by Jaffe et al. (1993). Patent citations are references to existing patents included in patent documents.
What does Regional Economic Integration Deliver?

In this article, I showed that recent RTAs are deeper than older ones. However, many recent RTAs are still restricted to cover the provisions that fall under the current mandate of the WTO discipline. Policy areas outside the WTO discipline are covered by a limited number of RTAs. Moreover, in terms of the depth of integration, the coverage of behind-the-border policies and other policies is relatively low, whereas the coverage of provisions on import tariffs and non-tariff barriers is high. In terms of the breadth of integration, on the other hand, many agreements cover services and investment categories, but the coverage of technology, labor, and non-economic policies is still low. Furthermore, I found that RTAs are heterogeneous in the depth and breadth according to types of RTAs, country pairs, and particular signatory countries, though the differences tend to be narrowing. Based on these observations, I conclude that deep integration in most regions is still halfway to its goal.

There is a growing literature on issues related to the impacts of deep regional economic integration. As reviewed in this article, a number of studies have shown that deep RTAs deliver many gains to trade in goods, FDI, GVCs, and R&D, whereas they may also deliver some losses. However, studies of deep RTAs are still scarce. Hence, we have not yet fully understood those gains and losses. In addition, the mechanism to deliver those gains and losses is still to be investigated.

Studies of deep RTAs have focused on their economic effects. However, people are also concerned about the non-economic benefits and losses. The citizens of the United Kingdom voted for Brexit despite the fact that it will end up with net economic losses.\(^{13}\) This implies that factors other than economic gains and losses may be important for people to decide their preferences for deep economic integration.

Rodrik (2011) provides an insightful explanation of this issue. He calls this the “political trilemma of the world economy.” His argument is as follows: (sovereign) \textit{nation states}, democratic politics, and hyper-globalization or deep international economic integration are mutually incompatible. Countries can choose at most two of these three options. Thus, if his argument is true, deep RTAs may be incompatible with national democracy. Only shallow RTAs may be compatible with the sovereignty of nations and democracy. This is an important proposition for both theoretical and empirical studies of deep RTAs. Related to this issue, Sampson (2017) argues that the “Leave” vote at the Brexit referendum in 2016 may be the assertion of national identity and that Brexit is “a democratic response to the erosion of British sovereignty caused by EU membership” (p. 180). Considering this important issue for deep regional integration, interdisciplinary research is needed to uncover the various (economic and non-economic) impacts of deep regional integration and understand people’s responses.

Finally, international cooperation is crucial for tackling pandemics. RTAs may even address policies for preventing diseases and protecting public health in member countries. Actually, “health” is one of the policy areas in the WTO-X group (see Table 1). However, only 17 RTAs out of 279 (that is, only about 6% of all RTAs) in the contents of RTA dataset, which entered into force by the end of 2015, include a health provision with legal enforceability.\(^{14}\) Therefore, in order to tackle COVID-19 pandemic, it may be a key to pro-

---

\(^{13}\) On the economic effects of Brexit, see, for example, Breinlich et al. (2018), Crowley et al. (2020), Dhingra et al. (2017), and Graziano et al. (2018).

\(^{14}\) For example, the Treaty on European Union includes articles on public health: “The Community shall contribute towards ensuring a high level of human health protection by encouraging cooperation
mote international cooperation through deep integration.

References

Antràs, P. and E. Helpman (2004), “Global sourcing,” Journal of Political Economy, vol. 112, no. 3, pp. 552–580.

Antràs, P. and E. Helpman (2008), Contractual frictions and global sourcing, in Helpman, E., Marin, D., and Verdier, T., eds., The Organization of Firms in a Global Economy. Harvard University Press, 9–54.

Baier, S.L., Y.V. Yotov, and T. Zylkin (2019), “On the widely differing effects of free trade agreements: Lessons from twenty years of trade integration,” Journal of International Economics, vol 116, p. 206–226.

Baldwin, R. (2011), 21st century regionalism: Filling the gap between 21st century trade and 20th century trade rules, World Trade Organization Staff Working Paper ERSD–2011–08.

Baldwin, R. (2016), The Great Convergence: Information Technology and the New Globalization, Harvard University Press.

Behar, A. and L. Cirera–i-Crivillé (2013), “Does it matter who you sign with? Comparing the impacts of North–South and South–South trade agreements on bilateral trade,” Review of International Economics, vol. 21, no. 4, pp. 765–782.

Boffā, M., M. Jansen, and O. Solleder (2018), “Do we need deeper trade agreements of GVCs or just a BIT?,” The World Economy, vol 42, pp. 1713–1739.

Breinlich, H., E. Leromain, D. Novy, T. Sampson, and A. Usman (2018), “The economic effects of Brexit: Evidence from the stock market,” Fiscal Studies, vol 39, no 4, pp. 581–623.

Cheong, J., D.W. Kwak, and K.K. Tang (2015), “Heterogeneous effects of preferential trade agreements: How does partner similarity matter?,” World Development, vol. 66, pp. 222–236.

Crowley, M.A., O. Exton, and L. Han (2020), “The looming threat of tariff hikes: Entry into exporting under trade agreement renegotiation,” AEA Papers and Proceedings, vol. 110, pp. 547–551.

Dhingra, S., H. Huang, G. Ottaviano, J. Paulo Pessoa, T. Sampson, and J. Van Reenen (2017), “The costs and benefits of leaving the EU: Trade effects,” Economic Policy, vol. 32, no. 92, pp. 651–705.

Dür, A., L. Baccini, and M. Elsig (2014), “The design of international trade agreements: Introducing a new dataset,” Review of International Organizations, vol. 9, pp. 353–375.

Graziano, A., K. Handley, and N. Limão (2018), Brexit uncertainty and trade disintegration, NBER Working Paper No. W25334.

Helpman, E., M.J. Melitz, and S.R. Yeaple (2004), “Export versus FDI with heterogeneous firms,” American Economic Review, vol. 94, no. 1, pp. 300–316.

Hofmann, C., A. Osnago, and M. Ruta (2019), “The content of preferential trade agreements,” World Trade Review, vol. 18, no. 3, pp. 365–398.

Horn, H., P.C. Mavroidis, and A. Sapir (2010), “Beyond the WTO? An anatomy of EU and US preferential trade agreements,” The World Economy, vol 33, no. 11, pp. 1565–1588.

Jaffe, A.B., M. Trajtenberg, and R. Henderson (1993), “Geographic localization of knowledge spillovers as evidenced by patent citations,” Quarterly Journal of Economics, vol. 108, pp. 577–598.

Jinji, N. and I. Kamata (2020), Do labor clauses in regional trade agreements reduce the trade creation effect?, Kyoto University, Graduate School of Economics Discussion Paper Series No. E-20–002.

Jinji, N., Y. Sawada, X. Zhang, and S. Haruna (2021), Deep regional trade agreements and cross-border technology transfer: Theory and evidence, Kyoto University, Graduate School of Economics Discussion...
Paper Series No. E-20-008.
Jinji, N., X. Zhang, and S. Haruna (2018), Does deep economic integration facilitate international research collaboration?, Unpublished manuscript.
Jinji, N., X. Zhang, and S. Haruna (2019), “Do deeper regional trade agreements enhance international technology spillovers?,” The World Economy, vol. 42, no. 8, pp. 2326–2363.
Laget, E., A. Osnago, N. Rocha, and M. Ruta (2020), “Deep trade agreements and global value chains,” Review of Industrial Organization, online first.
Lawrence, R.Z. (1996), Regionalism, Multilateralism, and Deeper Integration, Washington, DC: Brookings Institution.
Limão, N. (2016), Preferential trade agreements, in Bagwell, K. and Staiger, R.W., eds., Handbook of Commercial Policy, Vol. 1B. Amsterdam: Elsevier, 279–367.
Mattoo, A., A. Mulabdic, and M. Ruta (2017), Trade creation and trade diversion in deep agreements, World Bank Policy Research Working Paper No. 8206.
Orefice, G. and N. Rocha (2014), “Deep integration and production networks: An empirical analysis,” The World Economy, vol. 37, 106–136.
Osnago, A., N. Rocha, and M. Ruta (2016), “Do deep trade agreements boost vertical FDI?,” World Bank Economic Review, vol. 30, S119–S125.
Osnago, A., N. Rocha, and M. Ruta (2019), “Deep trade agreements and vertical FDI: The devil is in the details,” Canadian Journal of Economics, vol. 52, no. 4, pp. 1558–1599.
Rodrik, D. (2011), The Globalization Paradox: Democracy and the Future of the World Economy, WW Norton & Company.
Rodrik, D. (2018), “What do trade agreements really do?,” Journal of Economic Perspectives, vol. 32, no. 2, pp. 73–90.
Roy, J. (2010), “Do customs union members engage in more bilateral trade than free–trade agreement members?,” Review of International Economics, vol 18, no. 4, pp. 663–681.
Sampson, T. (2017), “Brexit: The economics of international disintegration,” Journal of Economic Perspectives, vol. 31, no. 4, pp. 163–184.
Vicard, V. (2009), “On trade creation and regional trade agreements: does depth matter?,” Review of World Economics, vol. 145, no. 2, pp. 167–187.
Vicard, V. (2011), “Determinants of successful regional trade agreements,” Economics Letters, vol 111, no. 3, pp. 188–190.
Viner, J. (1950), The Customs Union Issue, Carnegie Endowment for International Peace.