

ORIGINAL ARTICLE

Services Trade Policies and Economic Integration: New Evidence for Developing Countries

Bernard Hoekman^{1*}  and Ben Shepherd²

¹European University Institute, Italy and CEPR and ²Principal, Developing Trade Consultants, USA

*Email: bernard.hoekman@eui.eu; ben@developing-trade.com

(Received 11 December 2019; revised 19 August 2020; accepted 23 August 2020)

Abstract

This paper applies machine learning to recreate to a high degree of accuracy the OECD's Services Trade Restrictiveness Index (STRI) to provide quantitative evidence on the restrictiveness of services policies in 2016 for a sample of developing countries, using regulatory data collected by the World Bank and WTO. Resulting estimates are used to extend the OECD STRI approach to 23 additional countries, producing what we term a Services Policy Index (SPI). Converting the SPI to ad valorem equivalent terms shows that services policies are typically much more restrictive than tariffs on imports of goods, in particular in professional services and telecommunications. The SPI has strong explanatory power for bilateral trade in services at the sectoral level, as well as for aggregate goods and services trade.

JEL Codes: F13; F15; O24

Keywords: International trade; trade in services; machine learning; services policy; trade restrictiveness indicators

1. Introduction

Services play an important role in economic development. Because services account for a significant share of total output in even very poor countries, the operation of services sectors matters for overall economic performance. The importance of services for development is augmented because of their role as inputs into production for a broad cross-section of industries, including agriculture as well as manufacturing. The cost, quality, and variety of services available in an economy helps determine the productivity of 'downstream' sectors. Services also matter for the achievement of the sustainable development goals (SDGs); improving access to health, education, and finance or enhancing connectivity through investment in information and communications technologies and transport and logistics networks all involve services activities.¹

Restrictive trade and investment policies may impact negatively on firms using services as inputs, reduce the competitiveness of services exporters, and increase prices and/or lower the quality of services available to households.² Trade in services is like trade in goods in allowing specialization according to comparative advantage, inducing competitive pressures and knowledge spillovers, but differs in that often it is more dependent on cross-border movement of providers, whether legal entities (firms) or natural persons (services suppliers). A consequence is that trade in services involves a much broader range of policy instruments than trade in goods (Francois and Hoekman, 2010).³

¹See <https://sustainabledevelopment.un.org/topics/sustainabledevelopmentgoals> for more detail on SDG targets.

²See, e.g., Borchert et al. (2011), Fiorini and Hoekman (2018), and Helble and Shepherd (2019).

³Factor movement and cross-border movement of buyers/consumers is also salient for trade in goods but is generally much less likely to be a necessary condition for trade to be feasible.

Well-known data weaknesses hamper analysis of how policies towards imports and exports of services, foreign direct investment, and, more generally, regulation affects the operation of services sectors. Although data on services activities in developing economies has been improving, in part as the result of periodic firm-level surveys that have resulted in large panel datasets (e.g., the World Bank enterprise surveys), comparable information on external service-sector policies of developing countries is very limited, information on policies is often patchy at best, and time series data on relevant policy variables generally are not available on a cross-country, comparable basis.

This situation began to change in the late 2000s with a World Bank project to collect information on services trade and investment policies and to create services trade restrictiveness indicators (STRIs) that constitute a numerical summary of applied services policies believed to affect trade flows (Borchert et al., 2014). These STRIs in turn have been used to estimate sectoral ad valorem tariff equivalents for 103 countries (Jafari and Tarr, 2017). The OECD has gone further than the World Bank by compiling STRIs for its member countries as well as major emerging economies that span a broader range of policies and services sectors, including both discriminatory and regulatory measures. The OECD STRI is available on an annual basis starting in 2014 and covers 45 countries. The general characteristics of the two STRI projects are set out in Table 1.

A problem for applied policy research on developing country services trade policies is that the OECD STRI database covers only a small number of emerging countries, while the World Bank STRI data are only available periodically, with the latest STRI covering 2016 policies released in early 2020. As a result, extant empirical research on developing country services trade policies has largely been constrained to cross-section analysis, using increasingly outdated information.

The 2020 edition of the World Bank STRI is supported by a collaborative data collection effort with the WTO secretariat. A first result of this joint venture was the publication in November 2019 on the jointly managed Integrated Trade Intelligence Portal (I-TIP) website of a database of applied services trade policies for the year 2016. These data span many emerging and developing economies as well as OECD member countries. In January 2020, the World Bank released its own STRI based on these data.

In this paper, we utilize the World Bank–WTO information on 2016 services policies to generate new indicators of services policy restrictiveness in eight services sectors for 23 countries not included in the OECD STRI.⁴ The new data provide an opportunity to analyze services trade policies using information that post-dates the 2008 global financial crisis. In addition to developing a methodology to effectively extend coverage of the OECD index using the new data, we show that it is a strong predictor of bilateral services trade at the sectoral level.⁵ Finally, we take advantage of the recent release of the World Bank's STRI to compare performance of the various indicators in an empirical setting.

A challenge in generating indicators of services trade policy from information on applied measures is the need to appropriately weigh and aggregate policies on a sector-by-sector basis. A contribution of this paper is to apply a machine-learning algorithm to the policy data to construct indicators that are broadly consistent with the STRI methodology used by the OECD in that they correlate well with the OECD STRIs. Because the full detail of the methodology used to produce the OECD indices is proprietary and not published, it is not possible to simply apply the OECD methodology to generate STRIs that are strictly comparable to those reported in the OECD database. A similar problem arises in relation to the World Bank STRI released in

⁴The OECD produces STRIs for OECD member countries and nine (mostly large) emerging economies: Brazil, China, Colombia, Costa Rica, India, Indonesia, Malaysia, the Russian Federation, and South Africa. See <https://qdd.oecd.org/subject.aspx?Subject=063bee63-475f-427c-8b50-c19bffa7392d>. The additional country that is the focus of this paper is Rwanda for which data were produced by Shepherd et al. (2019b) with assistance from the OECD Secretariat. This brings the total to 24. See Appendix 1.

⁵As made clear below, the index itself is not based in any way on trade data.

Table 1. Comparison of key characteristics of OECD and World Bank STRIs

	OECD	World Bank
Years covered	2014, 2015, 2016, 2017, 2018, 2019.	2008, 2016.
Publication lag	< 1 year	4 years
Countries covered	46	68
Sectors covered	22	23
Data source	Collected directly from national laws and regulations	Sourced from OECD (43 countries) and law firm surveys (25 countries)
Regulatory data publicly available	Yes	Yes
Policy simulator	Yes	No
Algorithm code publicly available	No	No

2020. Although staff of both organizations have published papers outlining their methodologies, it is challenging for independent researchers to collect data for new countries and calculate compatible indices. Our approach provides a straightforward approach that researchers anywhere can reproduce to calculate their own STRIs based on newly collected policy data. An added benefit of our approach is that it highlights the individual policy measures that have the greatest explanatory power in terms of the sectoral STRIs and offers the perspective of streamlining data collection efforts in the future.

The plan of the paper is as follows. In Section 2, we discuss briefly the new data on 2016 services policies published by the WTO. Section 3 describes the methodology used to generate services policy indicators (SPIs) from this information. Section 4 validates the SPIs by assessing their ability to act as statistically significant predictors of trade flows using a standard structural gravity model of total trade and specific services sectors. Section 5 concludes.

2. New Data on Services Policies

In November 2019, the World Bank and WTO released an update to their jointly managed I-TIP platform containing extensive data on national services policies. In its raw state, the dataset includes 68 countries, 25 sectors, and three modes of supply: cross-border trade in services (Mode 1 in WTO speak), Mode 3 (establishment of a commercial presence in a foreign country – essentially foreign direct investment in a services sector), and Mode 4 (temporary cross-border movement of services suppliers). The data exclude Mode 2, where trade occurs through movement of consumers to a foreign country (e.g., tourism) as this is generally unrestricted.

The dataset pertains to policies observed in 2016 that potentially affect services trade. It has nearly a quarter of a million observations (244,949), distinguishing up to 445 different measures, both sector specific and horizontal. If attention is restricted to countries and sectors for which information is reported fully at the level of these individual measures, the coverage falls to 68 countries and 24 sectors.⁶ I-TIP data are freely downloadable from the WTO website. Borchert et al. (2019) discuss database creation, including sourcing and access. The source for 45 of the 68 countries is the OECD STRI database, so that I-TIP adds information on 23 countries not

⁶Some of the measures are coded for only a handful of countries, precluding use in empirical analysis in a cross-country setting. As it is important for empirical analysis to have data availability across all relevant data points, we limit consideration to the countries and sectors we have identified as satisfying that criterion.

Table 2. Classification of World Bank/WTO services policy data

A	<i>Conditions on market entry</i>
1	Forms of entry (including foreign equity limits)
2	Quantitative and administrative conditions
3	Conditions on licensing/qualifications relating to market entry
4	Other conditions on market entry
B	<i>Conditions on operations</i>
1	Conditions on supply of services
2	Conditions on service supplier
3	Conditions on government procurement
4	Other conditions on operations
C	<i>Measures affecting competition</i>
1	Conditions on conduct by firms
2	Governmental rights/prerogatives (including public ownership)
3	Other measures affecting competition
D	<i>Regulatory environment and administrative procedures</i>
1	Regulatory transparency (including licensing)
2	Nature of regulatory authority (measures related to nature of regulator)
3	International standards
4	Conditions related to administrative procedures
5	Other regulatory environment and administrative procedures
E	<i>Miscellaneous measures</i>

Source: Borchert et al. (2019).

covered by the OECD (Appendix Table 1 lists the countries). As with the 2008 iteration of the World Bank STRI, questionnaires administered to law firms in the countries of interest generated the raw data, treated by the World Bank and WTO team to ensure consistency and correctness. Table 2, taken directly from Borchert et al. (2019), lists the general categories of measures included in the database.

3. Constructing an Index of Services Policies from I-TIP Data

There are two key analytical decisions in designing an STRI given the choice to collect data on particular measures: weighting the measures and aggregating them into an index. The first problem can be solved through application of purely statistical methods, such as factor analysis, or by using external expert judgment, as in the OECD STRI, which is based on a weighting and aggregation system driven by expert input (Grosso et al., 2015). Aggregation is then by weighted sum. The World Bank STRI, by contrast, appears to use the judgment of World Bank analysts to motivate the mapping of measures to quantitative scores. Then aggregation uses a constant elasticity of substitution function with parameters chosen again by the judgment of World Bank analysts.

As noted above, the selection of I-TIP data we use span 455 individual policy measures in 68 countries and 24 sectors. The challenge is to produce an overall index of services policy by sector using those data. We use the OECD index as our basis for this problem because it appears to us to

represent the best currently available STRI.⁷ It is quite possible that a superior index will be designed in the future, but, at the present time, we do not believe there is a better benchmark than the OECD STRI. There is an active research program based on it, showing that the index is robustly linked with trade in services (e.g., Nordås and Rouzet, 2017) and investigating questions such as the extent and effects of regulatory heterogeneity (Nordås, 2016, 2018) and the services content of regional integration in the EU (Benz and Gonzalez, 2019).⁸ Moreover, we are interested in showing the potential for commonplace statistical techniques to extend coverage of an index to new countries as additional data become available.

The first problem we address in this paper is how to extend the OECD STRI to the countries included in I-TIP but not in the OECD database. The purpose of undertaking this exercise is to show that commonplace quantitative methodologies can reproduce the STRI very well, even without the detailed information contained in the aggregation algorithm, which has not been made public. Our analysis illustrates it is possible to obtain a comparable index across all 68 countries quickly and at very low cost, without the need to develop a new methodology.

Equivalently, our problem can be seen as one of dimensionality reduction. We start with up to 445 individual policy measures per country in I-TIP and attempt to reduce those measures to one score per country–sector pair. Our objective in reducing dimensionality is to reproduce the OECD STRI as closely as possible.

This class of problem is well suited to a basic machine learning application. We construct a dataset containing OECD STRIs by sector and all horizontal and sector specific policy measures from I-TIP for all 68 countries for which full data are available. For the analysis to be feasible, we limit consideration to those sectors that correspond well between the two databases, taking simple averages of measures where necessary. This approach is made necessary by the different levels of aggregation in the data: subsectors in I-TIP, and more aggregate sectors in the publicly available OECD data. The result is to reduce the number of sectors we can work with easily to eight: accounting, legal, commercial banking, insurance, air transport, road freight transport, distribution, and telecom. We believe these sectors represent a large share of services activity in most countries. Although we lose some of the nuance in the I-TIP data – which distinguishes sectors at a micro level, such as insurance versus reinsurance, or air passenger transport versus air cargo transport – we believe this approach is justifiable given our overall objectives as set out above.

We split the sample into three groups. We randomly assign 75% of observations for which there is an OECD STRI to a training subsample, with the remaining 25% assigned to a prediction subsample. Finally, those countries and sectors where no OECD STRI is available are assigned to an out of sample prediction subsample. The basic empirical strategy is to apply machine learning models to the training subsample only, with the aim of using the disaggregated policy data to predict an index score by sector, which can then be compared with the OECD STRI. By having the model produce indices for the prediction subsample, we can assess how it performs on data that were not used for training. Once an appropriate model has been selected, it can be used to make out of sample predictions for countries not in the OECD data.

⁷Although the OECD STRI is the best available index, not least because the organization undertook an extensive attempt to obtain weights by expert judgement, these experts defined the various weights used to define the index on the basis of what they presumed is the best relative importance of each policy measure. As noted by a reviewer, to determine this relative importance objectively, ideally one should have weights *ex ante*, based on trade data. The current state of services trade data availability does not permit this to be done.

⁸The body of evidence using the World Bank STRI is smaller, likely reflecting the intermittent nature of the exercise which has historically limited researchers to cross-sectional analysis. See e.g., Borchert et al. (2014), Hoekman and Shepherd (2017), Beverelli et al. (2017), and Su et al. (2019) for analyses using the World Bank STRI.

3.1 Developing Services Policy Indices with Simple Machine Learning

Machine learning refers to a family of techniques commonly used in computer science and data science, whereby statistical algorithms are used to find patterns in data, and to make predictions based on those patterns. These types of approaches power many familiar consumer interfaces, such as movie suggestions on Netflix, or predictive text in Google searches or Gmail emails. To date, machine learning has been applied infrequently in economics, but Athey and Imbens (2019) identify features of machine learning techniques that can potentially be of use to applied economists. An important example of such an approach is Athey et al. (2017), who use machine learning to assist in developing a causal inference framework. The remainder of this section sets out our general approach. Readers who do not require the technical background can skip to the next section.

We apply a type of machine learning model called an elastic net as a prediction tool. The model itself can be understood as similar to the standard regression framework familiar to economists, with independent variables (inputs) and a dependent variable (output), with the relationship between the two summarized in a set of coefficients that are estimated by solving an optimization problem. In this case, the input is the set of policy measures from I-TIP, and the output is the OECD STRI. Model selection and performance assessment proceed by comparing the model’s predictions of the STRI with what is observed. A good model is one that uses the inputs to reproduce the OECD’s STRI to a high degree of accuracy, and which exhibits similar accuracy for the subsample of data used to train the model, and the remainder of the sample that was not used in training. A subsidiary consideration is model parsimony, namely the ability to achieve strong predictive performance while using a relatively small number of variables. The machine learning techniques we use automatically drop policy measures (inputs) with little explanatory power for the OECD STRI (the output) and identify a set of policy measures that have the greatest explanatory power, typically much smaller than the starting data set. We discuss the policy implications of this feature of the algorithms below.

In technical terms, the elastic net solves the following problem, where $\hat{\beta}$ is the vector of parameters of interest:

$$\hat{\beta} = \operatorname{argmin}_{\beta} \left\{ \frac{1}{2n} \sum_{i=1}^n (y_i - x_i\beta')^2 + \lambda \left[\alpha \sum_{j=1}^p |\beta_j| + \frac{1-\alpha}{2} \sum_{j=1}^p \beta_j^2 \right] \right\}$$

The first term is the standard ordinary least squares (OLS) loss function. λ is a penalty term that shrinks parameter estimates towards zero in two ways, with a higher parameter resulting in greater shrinkage. The first term in square brackets penalizes coefficients that are large in absolute value, while the second performs shrinkage based on the square of the parameter value. With $\lambda = 0$, the elastic net collapses to standard OLS. With nonzero λ and $\alpha = 1$, it is the least absolute shrinkage and selection operator (LASSO), while with $\alpha = 0$, it is ridge regression. The essence of the procedure is that λ is iterated for given values of α , with zero coefficients dropped from the model progressively due to the shrinkage effect. Iteration continues until a model is selected based on its cross-validation performance, i.e. the ability of a model estimated on the training subsample only to produce close estimates of the values in the prediction subsample. By proceeding in this way, we can identify a subset of variables that have the best explanatory power in terms of the observed OECD STRI, and then use the estimated values from the elastic net regression to predict values out of sample, where no OECD STRI exists.

The elastic net is well suited to prediction problems with large numbers of potential predictors, even exceeding the number of observations, and deals well with situations where they are closely correlated. These characteristics are unlike OLS, where the number of potential predictors cannot be greater than the number of observations. More fundamentally, the objective of OLS is typically inference, whereas the objective of the elastic net is prediction. As a result, we do not report

standard errors or statistical significance, which are typically important outputs of an OLS model. Instead, we report cross-validation performance, namely the ability of the model to make predictions that closely track observed values in the subsample not used to train the model in the first place.

To power the tool, we construct a set of explanatory variables that is all sectoral policy measures from I-TIP, all horizontal policy measures from I-TIP, and a set of sector dummies. We then also create interactions to allow for nonlinear effects and dependencies. Specifically, we interact all policy measures with all other policy measures, and we create a triple interaction between all horizontal policy measures, all sector specific policy measures, and the sectoral dummies. The I-TIP dataset contains missing entries for many response variables, presumably because they are believed to only be relevant to certain sectors. To facilitate the empirical analysis, we therefore code these missing values as zero, which means that they do not have any restrictive impact on trade in sectors where World Bank and WTO analysts have made an a priori determination of no effect.

Proceeding in this way gives a dataset of 544 observations, which is eight sectors for 68 countries. By interacting all the potential explanatory variables, as set out above, we have 16,974 variables. Many of those variables are constant within subsamples, often zero, and so are automatically dropped from the model. In practice, the elastic net works with a starting set of 1,606 variables. A standard regression technique like OLS cannot handle this problem given the number of observations, but the elastic net can, because the optimization problem has kinks due to the absolute value and square terms. Since OLS is unavailable we use two other dimension reduction techniques on the sectoral and horizontal measures to give a point of comparison but ignoring interaction terms: principal factor analysis and a simple mean. In both cases, we apply the relevant measures to the disaggregated policy data at the level of sectors, to produce a single index that summarizes, according to the assumptions of each, the variation in the disaggregated policies. As a robustness check on our machine learning models, we also set $\alpha = 1$, which yields LASSO estimates, and $\alpha = 0$, which yields ridge estimates.

Given that the problem in this case is prediction, not inference, we do not report coefficient estimates. For the training sample (272 observations), the elastic net retains 59 variables, a mix of measures in levels and interactions, and selects $\alpha = 0.25$. The LASSO retains 55 variables, while the ridge estimator retains the full set of informative variables, namely 1,606. Table 3 summarizes the performance of the three machine learning methods, looking separately at the training and prediction subsamples.

The three methods perform quite similarly on the training subsample: model fit is tight considering the relatively small amount of information used. The mean value of the OECD STRI is 0.279, so a mean squared error of only 0.005 using the elastic net indicates that model fit is good. Comparing the two parts of Table 3 shows that of the three machine learning methods, the elastic net has the best performance: R^2 is highest both on the training and prediction subsamples. We therefore prefer the elastic net's predictions, but we note that it is relatively close in performance to the other two models.

Table 4 reports the correlations at the sectoral level among the various measures computed as described above. The elastic net again is the strongest performer on this overall criterion, although the other two machine learning methods also perform well. The comparator indices, constructed using principal factor analysis and a simple mean, have a negative correlation with the OECD index, and thus represent a radically different way of summarizing the data. The evidence in Table 4 suggests that the OECD approach to weighting and aggregating measures results in an output that is substantially different from what can be obtained by naïve methods that focus only on the characteristics of the data, without using external knowledge on the effects of particular policies. It is an important finding that the OECD index, and by extension, ours, does not correlate strongly or even positively with naïve measures. But our three simple machine learning applications, using limited data, do a remarkable job of reproducing the OECD index. Moreover,

Table 3. Output from Elastic Net, LASSO, and ridge applications to OECD STRIs using I-TIP data in levels and interactions

	Mean Squared Error	R-Squared	Observations
Training			
Elastic Net	0.005	0.784	272
LASSO	0.008	0.683	272
Ridge	0.009	0.594	272
Prediction			
Elastic Net	0.007	0.739	91
LASSO	0.009	0.674	91
Ridge	0.012	0.527	91

Table 4. Correlation between OECD STRI and alternative services policy index (SPI) estimates (common sample)

	OECD STRI	Elastic Net SPI	LASSO SPI	Ridge SPI	Principal Factors SPI	Simple Mean SPI	World Bank STRI
OECD STRI	1.000						
Elastic Net SPI	0.888	1.000					
LASSO SPI	0.851	0.983	1.000				
Ridge SPI	0.816	0.909	0.869	1.000			
Principal Factors SPI	-0.266	-0.327	-0.350	-0.430	1.000		
Simple Mean SPI	-0.357	-0.415	-0.416	-0.536	0.780	1.000	
World Bank STRI	0.822	0.800	0.786	0.768	-0.208	-0.361	1.000

our preferred method, the elastic net, produces predicted values that lie exclusively between zero and unity, as does the original OECD index. The alternative approaches do not have this property, nor would a simple OLS regression model. Given that the indices only have an ordinal interpretation, it is useful to investigate the rank correlation as well. We find that the OECD STRI and our SPI have a rank correlation coefficient of 0.788, which is again evidence of a strong correspondence between the two measures.

Figure 1 shows the correlation in a common sample of elastic net predictions for all countries in the I-TIP sample, i.e. those already covered by the OECD and those that are new, which we name the Services Policies Index (SPI), and the OECD STRI at the sector level. The association with the OECD index is not perfect, as would be expected with any statistical approach to reproduction of an existing index, but the figure shows that our SPI fits the original data well, which gives us confidence that out of sample estimates for the countries not in the OECD database should perform well, in particular given the similarity of the R^2 measures for the training and prediction sub-samples, as noted above. Summary statistics show that the new countries added to the sample are more restrictive than those in the original OECD sample, with an average STRI across sectors of 0.333 compared with 0.281; however, ranges in both cases are wide: the new countries have scores ranging from 0.096 to 0.872, compared with 0.157 to 0.960 for those in the OECD sample.

To avoid terminological confusion in what follows we refer consistently to the OECD STRI as the STRI. Our constructed indices based on I-TIP data are referred to as Services Policy Indices (SPIs). The difference in terminology highlights that we are simply mimicking the OECD's

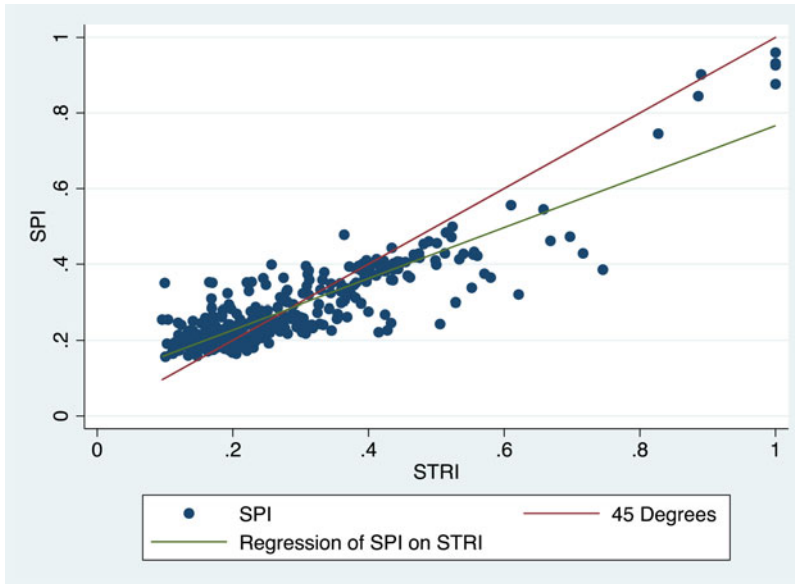


Figure 1. Correlation between the STRI and SPI, sector level (common sample)

original approach using a broader dataset. Ownership of the full methodology used to produce the OECD's indices lies with that organization, and we use a commonplace data-driven technique to extend database coverage.

4. Validating the SPI with Trade Data

We have already shown that our SPI closely mirrors the OECD's STRI, in our view currently the best available benchmark for measuring applied services policies. An important additional step in validating the SPIs is demonstrating their ability to act as statistically significant predictors of trade flows. Thus far, we have not used any information on trade flows to produce the SPI, but only the weights implied by the OECD's expert judgment system, which we have mimicked using machine learning. This second step is therefore quite independent of the first, and is designed to show that our index has explanatory power for bilateral trade, which we believe is an important criterion for any index that seeks to measure services trade policies.

We estimate a standard structural gravity model of services trade at the sectoral level in line with current best practice, as embodied by Anderson et al. (2018). Estimation is by Poisson Pseudo Maximum Likelihood (PPML), which means that estimates are robust to heteroskedasticity, take account of zero flows, and produce fixed effects (by exporter and by importer) that correspond exactly to the quantities prescribed by theory in Anderson and Van Wincoop (2004)-type models (Fally, 2015).

The gravity model takes the following form, considering a single year and single sector cross-section only:

$$X_{ij} = F_i F_j t_{ij}^{-\theta} e_{ij} \quad (1)$$

where: X_{ij} is exports from country i to country j ; the F terms are exporter and importer fixed effects; t_{ij} is bilateral trade costs; θ is a parameter capturing the sensitivity of demand to cost;

and e_{ij} is an error term satisfying standard assumptions.⁹ Trade costs t are specified in the usual iceberg form. These costs are unobserved but can be specified in terms of observable proxies. For present purposes, we include standard gravity model controls based on geography and history, and an indicator of service sector restrictiveness (STRI for presentational purposes), as well as an interaction between the STRI and a dummy for countries that are members of an Economic Integration Agreement (EIA) under Article V of the GATS, the services equivalent of a PTA for goods. Formally:

$$-\theta \log t_{ij} = b_1 \text{STRI}_{ij} * \text{intl}_{ij} + b_2 \text{EIA}_{ij} + b_3 \text{STRI}_{ij} * \text{intl}_{ij} * \text{EIA}_{ij} + b_4 \log(\text{distance}_{ij}) + b_5 \text{contiguous}_{ij} + b_6 \text{colony}_{ij} + b_7 \text{common language}_{ij} + b_8 \text{common colonizer}_{ij} + b_9 \text{same country}_{ij} + \text{intl}_{ij} \quad (2)$$

Table 5 provides variable definitions and sources, along with those for equation (1). Apart from trade flows, the data sources are largely standard. Equation (1) should in principle cover all directions of trade, i.e. including trade from country i to country i , or intra-national trade. Inclusion of intra-national trade data is crucial for PPML to produce theory-consistent fixed effects estimates (Fally, 2015). International trade data do not include this term, so we use the Eora multi-region input–output table to do the job.¹⁰ Eora covers 183 countries and 26 sectors through a single harmonized input–output table. We use data for 2015 only, the latest available year, corresponding most closely to the year of our SPI data (2016).

In sector by sector regressions, the STRI is a country-specific variable and so is collinear with the importer fixed effects. To achieve identification, we follow Heid et al. (2017) in interacting it with *intl*, a dummy equal to unity for trade between different countries, and zero for internal trade. The reason for this approach is that the STRI has a differential impact on domestic and foreign service providers, which is a key aspect of the rationale for the index in the first place. We divide the World Bank STRI by 100, so that it is on the same scale as the other policy indicators.

As noted above, our SPI data start from 24 sectors defined in the World Bank/WTO dataset, which we concord to eight sectors in the OECD STRI classification. We then further concord those data to four Eora sectors by taking simple averages of the relevant indices: distribution, finance and business services, telecom, and transport. It is not possible to estimate gravity models at a more detailed level as the Eora database in harmonized form is necessarily highly aggregated.

A second point that requires explanation is the interaction term between services policies and EIA membership. The services policies in I-TIP apply on a most favored nation (non-preferential) basis, which is why we map them to MFN policies from the OECD data. The OECD has collected preferential data for services trade within the EU, but there is no systematic dataset covering preferential services policies around the world. However, many countries are members of trade agreements that hope to boost exports between their members, relative to the MFN benchmark. By interacting MFN policies with a dummy for joint EIA membership, we seek to capture that effect. An EIA in this context is effectively a trade agreement for services, under Article V of the GATS. Our expectation is that the coefficient on MFN policies will be negative (trade reducing), while the coefficient on the interaction term will be positive (showing that trade reduction is attenuated by regional integration). Benz and Gonzalez (2019) show conclusively in the case of the EU that intra-bloc services policies are far more liberal than those pertaining to non-EU countries.

⁹Numerous theoretical frameworks are consistent with this model, including as the Armington-type model of Anderson and Van Wincoop (2003), the Ricardian model of Eaton and Kortum (2002), and the heterogeneous firms model of Chaney (2008). Arkolakis et al. (2012) and Costinot and Rodriguez-Clare (2014) show that a wide class of quantitative trade models, including the canonical ones just cited, have the same macro-level implications for the relationship between trade flows and trade costs even though their micro-level predictions are quite different.

¹⁰See <https://worldmrio.com/>.

Table 5. Variables, definitions, and sources

Variable	Definition	Source
Colony	Dummy variable equal to one if one country in a pair was in a colonial relationship with the other.	CEPII.
Common colonizer	Dummy variable equal to one if the two countries were colonized by the same power.	CEPII
Common language	Dummy variable equal to one if both countries in a pair have a language in common, spoken by at least 9% of the population.	CEPII.
Contiguous	Dummy variable equal to one if the two countries share a common land border.	CEPII.
EIA	Dummy variable equal to one if the two countries are members of the same Economic Integration Agreement under Article V of the GATS.	Egger and Larch (2008).
Exports	Gross exports from country <i>i</i> to country <i>j</i> in sector <i>s</i> (2015).	Eora.
Intl	Dummy variable equal to one if country <i>i</i> and country <i>j</i> are different.	Authors.
SPI	Services Policies Index (Elastic Net, Lasso, Principal Factors, and Simple Mean).	Authors.
Log(Distance)	Logarithm of distance between country <i>i</i> and country <i>j</i> .	CEPII.
Log(Tariff)	Logarithm of 1 + applied tariff rate.	TRAINS
PTA	Dummy variable equal to one if country <i>i</i> and country <i>j</i> are part of the same preferential trade agreement in 2015.	Egger and Larch (2008).
Same Country	Dummy variable equal to one if the two countries were ever part of the same country.	CEPII.
STRI	OECD Services Trade Restrictiveness Index.	OECD
STRI WB	World Bank Services Trade Restrictiveness Index	World Bank

Table 6 reports gravity model regression results for the distribution sector. Column 1 includes the OECD STRI, and as expected, the policy variable has a negative coefficient, while the interaction term with EIA membership has a positive one, with both estimates statistically significant at the 10% level. The same is true for the World Bank STRI, which is included in column 7. The baseline data therefore support the view above that the measures captured by the STRI tend to restrict trade, in line with Nordås and Rouzet (2017), with that effect attenuated by joint membership of a trade agreement covering services. The same patterns of signs and magnitudes applies for four SPIs, elastic net, LASSO, ridge, and principal factors. The simple mean has no statistically significant coefficients. We therefore conclude that the most naïve of our testbed of SPI measures does not have significant predictive value for trade, but that other measures that attempt to summarize the available data more systematically do have such power.

Table 7 repeats the exercise for financial and business services. Results are similar to those for distribution. The elastic net, LASSO, and ridge SPIs perform somewhat better than the OECD STRI and on par with the World Bank STRI, in that the levels term and the interaction term both have coefficients with the expected signs and magnitudes, and are statistically significant at the 5% level or better. This is largely due to increased sample size for the SPIs; additional results, available on request, show that results for the SPI are less precisely estimated on the smaller sample for which the OECD STRI and the SPI are both available. Column 1 contains data on 183 exporters and 45 importers, while the following five columns all use 183 exporters and 68 importers. The principal factors SPI does not have any statistically significant coefficients, while the simple mean SPI has a negative and 1% statistically significant coefficient in levels,

Table 6. Gravity models for distribution services using different measures of services policies

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	STRI	Elastic Net	Lasso	Ridge	PF	Mean	WB
STRI*Intl	-5.617 *	-5.960 ***	-5.694 **	-7.678 ***	-1.755 **	0.532	-9.418***
	(3.076)	(2.263)	(2.388)	(2.393)	(0.730)	(0.482)	(2.263)
STRI*Intl*EIA	3.735 *	5.218 **	5.294 **	8.141 ***	2.633 ***	-0.428	0.055 ***
	(2.154)	(2.285)	(2.676)	(2.317)	(0.724)	(0.408)	(0.021)
EIA	-0.424	-0.607	-0.680	-1.389 ***	0.109	1.469	-1.420 *
	(0.389)	(0.479)	(0.602)	(0.530)	(0.150)	(0.939)	(0.743)
Log(Distance)	-0.328 ***	-0.333 ***	-0.335 ***	-0.320 ***	-0.308 ***	-0.343 ***	-0.345 ***
	(0.085)	(0.086)	(0.085)	(0.087)	(0.087)	(0.084)	(0.081)
Contiguous	0.675 ***	0.356	0.377	0.369	0.362	0.382	0.409
	(0.215)	(0.250)	(0.252)	(0.255)	(0.255)	(0.255)	(0.262)
Colony	0.316	0.329	0.370 *	0.299	0.455 **	0.357 *	0.342 *
	(0.245)	(0.207)	(0.199)	(0.210)	(0.192)	(0.213)	(0.188)
Common Language	0.112	0.393 **	0.388 **	0.389 **	0.396 **	0.386 **	0.351 **
	(0.176)	(0.177)	(0.179)	(0.172)	(0.176)	(0.186)	(0.168)
Common Colonizer	0.374	-0.102	-0.209	-0.095	-0.301	-0.235	-0.171
	(0.610)	(0.367)	(0.364)	(0.379)	(0.377)	(0.368)	(0.397)
Same Country	0.388	0.967 ***	1.013 ***	1.120 ***	1.054 ***	0.956 ***	1.451 ***
	(0.379)	(0.296)	(0.290)	(0.304)	(0.305)	(0.294)	(0.320)
Intl	-5.734 ***	-5.662 ***	-5.664 ***	-5.206 ***	-6.763 ***	-8.164 ***	-3.678 ***
	(0.688)	(0.574)	(0.628)	(0.612)	(0.277)	(1.063)	(0.845)
Constant	8.230 ***	8.074 ***	8.083 ***	7.990 ***	7.923 ***	8.135 ***	8.161 ***
	(0.528)	(0.523)	(0.518)	(0.530)	(0.529)	(0.508)	(0.490)
Observations	8418	12444	12444	12444	12444	12444	12261
R2	0.986	0.983	0.983	0.983	0.983	0.982	0.983
Importer fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exporter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: All models are estimated by PPML Robust standard errors adjusted for clustering by country pair in parentheses below parameter estimates. Statistical significance is indicated as follows: * (10%), ** (5%), and *** (1%).

but a statistically insignificant coefficient for the interaction term. The most naïve measures of services policies again have at best limited explanatory power, in contrast to more sophisticated measures like the STRI and the SPIs.

Table 8 reports results for telecom services. The pattern of findings is again quite similar: the STRIs, as well as the elastic net, LASSO, and ridge SPIs, all have explanatory power for bilateral trade flows in this sector, although none of the interaction terms except for the LASSO model has a statistically significant coefficient, which suggests that regional integration may not be a strong force for global trade in this sector. By contrast, the principal factors and simple mean SPIs have positive and 1% statistically significant coefficients, which is contrary to expectations.

Finally, Table 9 presents results for the transport sector. Both STRIs, elastic net SPI, and ridge SPI all have 5% statistically significant coefficients or better in levels and on the interaction term. By contrast, the principal factors SPI and the simple mean SPI do not have any

Table 7. Gravity models for finance and business services, STRI and SPIs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	STRI	Elastic Net	Lasso	Ridge	PF	Mean	WB
STRI*Intl	-1.620 ** (0.639)	-3.359 *** (1.002)	-3.819 *** (1.174)	-5.154 *** (1.630)	0.516 (1.058)	-3.459 *** (0.771)	-2.116 ** (0.931)
STRI*Intl*EIA	1.078 (0.689)	2.776 *** (1.034)	3.020 ** (1.195)	4.807 ** (1.871)	1.166 (0.984)	0.040 (0.670)	0.029 *** (0.009)
EIA	0.095 (0.216)	-0.126 (0.307)	-0.184 (0.354)	-0.726 (0.540)	0.694 *** (0.112)	0.724 (0.565)	-0.680 (0.442)
Log(Distance)	-0.470 *** (0.061)	-0.365 *** (0.070)	-0.364 *** (0.069)	-0.372 *** (0.068)	-0.357 *** (0.068)	-0.396 *** (0.066)	-0.386 *** (0.069)
Contiguous	0.421 *** (0.154)	0.552 *** (0.168)	0.553 *** (0.168)	0.553 *** (0.166)	0.597 *** (0.165)	0.525 *** (0.171)	0.542 *** (0.166)
Colony	0.163 (0.167)	0.211 (0.155)	0.220 (0.154)	0.196 (0.155)	0.289 ** (0.145)	0.350 ** (0.146)	0.206 (0.157)
Common Language	0.432 *** (0.103)	0.527 *** (0.106)	0.532 *** (0.107)	0.523 *** (0.106)	0.526 *** (0.107)	0.520 *** (0.105)	0.546 *** (0.105)
Common Colonizer	0.309 (0.345)	0.609 (0.660)	0.596 (0.663)	0.540 (0.654)	0.441 (0.672)	0.062 (0.601)	0.521 (0.667)
Same Country	0.248 (0.275)	0.321 (0.228)	0.320 (0.231)	0.357 (0.229)	0.226 (0.218)	0.183 (0.215)	0.267 (0.207)
Intl	-5.253 *** (0.260)	-5.267 *** (0.307)	-5.155 *** (0.336)	-4.719 *** (0.483)	-6.336 *** (0.215)	-3.493 *** (0.639)	-5.206 *** (0.423)
Constant	10.268 *** (0.379)	9.448 *** (0.429)	9.442 *** (0.428)	9.487 *** (0.419)	9.401 *** (0.418)	9.641 *** (0.404)	9.587 *** (0.424)
Observations	8418	12444	12444	12444	12444	12444	12261
R2	0.991	0.989	0.989	0.989	0.989	0.989	0.989
Importer fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exporter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: All models are estimated by PPML. Robust standard errors adjusted for clustering by country pair are in parentheses below parameter estimates. Statistical significance: * (10%), ** (5%), and *** (1%).

statistically significant coefficients. Results for this sector therefore accord well with those from the other sectors.

As is typical for gravity models that include intranational trade data, there is evidence of home bias in all regressions: the intl dummy is consistently negative and statistically significant. Also of interest is the finding that, in some specifications, the effect of an EIA effectively neutralizes the negative trade impact of the SPI or STRI. While deserving of further research, this result suggests that some regional agreements do a good job of facilitating services trade beyond the level that would be expected based on MFN policy settings.

Taken together, these results indicate that the OECD STRI has much greater explanatory power for bilateral trade flows in services than naïve measures like a principal factor or simple mean. The same is true of the World Bank STRI, which is unsurprising given how closely correlated it is with the OECD index ($\rho = 0.8$). Moreover, our three SPIs generally exhibit very similar performance to the OECD STRI, albeit with a substantially larger sample

Table 8. Gravity models for telecom services using different measures of services policies

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	STRI	Elastic Net	Lasso	Ridge	PF	Mean	WB
STRI*Intl	-4.389 *** (0.709)	-10.117 *** (1.494)	-11.247 *** (1.981)	-12.934 *** (2.160)	4.059 *** (0.645)	0.817 *** (0.156)	-5.953 *** (1.107)
STRI*Intl*EIA	-0.307 (0.738)	2.957 (1.969)	4.577 * (2.703)	0.647 (2.099)	-0.030 (0.593)	-0.136 (0.171)	0.000 (0.010)
EIA	0.183 (0.201)	-0.357 (0.463)	-0.684 (0.635)	0.095 (0.477)	0.465 (0.582)	0.790 * (0.445)	0.389 (0.400)
Log(distance)	-0.604 *** (0.058)	-0.479 *** (0.071)	-0.495 *** (0.073)	-0.486 *** (0.067)	-0.511 *** (0.065)	-0.514 *** (0.069)	-0.554 *** (0.067)
Contiguous	0.529 *** (0.151)	0.668 *** (0.164)	0.668 *** (0.174)	0.670 *** (0.153)	0.754 *** (0.174)	0.734 *** (0.171)	0.652 *** (0.172)
Colony	0.018 (0.188)	0.071 (0.162)	0.148 (0.171)	-0.038 (0.164)	0.245 (0.185)	0.218 (0.151)	0.165 (0.184)
Common language	0.304 *** (0.103)	0.447 *** (0.107)	0.368 *** (0.107)	0.533 *** (0.116)	0.335 *** (0.114)	0.298 *** (0.106)	0.304 *** (0.103)
Common colonizer	0.166 (0.234)	0.556 (0.526)	0.442 (0.527)	0.466 (0.500)	0.219 (0.509)	0.294 (0.548)	0.237 (0.514)
Same country	0.226 (0.368)	0.185 (0.227)	0.190 (0.239)	0.125 (0.221)	0.182 (0.270)	0.129 (0.270)	0.179 (0.249)
Intl	-4.142 *** (0.259)	-3.236 *** (0.423)	-2.955 *** (0.527)	-2.630 *** (0.473)	-9.568 *** (0.697)	-7.530 *** (0.411)	-3.218 *** (0.445)
Constant	9.106 *** (0.366)	8.113 *** (0.438)	8.211 *** (0.452)	8.151 *** (0.417)	8.309 *** (0.404)	8.328 *** (0.427)	8.591 *** (0.417)
Observations	8235	12444	12444	12444	12444	12444	12261.000
R2	0.975	0.972	0.972	0.972	0.972	0.971	0.972
Importer fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exporter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: All models are estimated by PPML. Robust standard errors adjusted for clustering by country pair in parentheses below parameter estimates. Statistical significance: * (10%), ** (5%), and *** (1%).

due to greater importer coverage. The difference in number of observations is just over 50%, so there are clear advantages to these extended measures based on data collected by the World Bank/WTO but aggregated into indices based on our machine learning-based reproduction of the OECD's approach. Given the strong and consistent explanatory power of the STRI and its derivative SPIs, the bar for producing a 'better' indicator of services trade restrictions is very high. In the absence of substantial additional benefits, it is far from obvious that further work in this area – in the sense of changing weights or adopting different aggregation schemes – passes a cost-benefit test, given the substantial time and resources that need to be devoted to dealing with the problems of weighting and aggregation discussed above.

Table 9. Gravity models for transport services using different measures of services policies

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	STRI	Elastic Net	Lasso	Ridge	PF	Mean	WB
STRI*Intl	-8.360 *** (1.643)	-4.375 * (2.389)	-1.168 (2.602)	-8.532 * (4.968)	-1.581 (0.967)	0.217 (0.366)	-5.419 *** (1.296)
STRI*Intl*EIA	4.693 *** (1.377)	4.624 ** (2.159)	4.913 ** (2.439)	8.544 ** (3.680)	1.229 (0.876)	-0.195 (0.292)	0.031 *** (0.010)
EIA	-1.139 ** (0.483)	-0.741 (0.692)	-0.785 (0.746)	-1.951 * (1.165)	1.439 ** (0.594)	0.898 *** (0.272)	-0.652 (0.486)
Log(distance)	-0.446 *** (0.066)	-0.320 *** (0.057)	-0.328 *** (0.058)	-0.320 *** (0.058)	-0.329 *** (0.057)	-0.323 *** (0.056)	-0.352 *** (0.056)
Contiguous	0.400 *** (0.143)	0.430 *** (0.151)	0.420 *** (0.155)	0.442 *** (0.152)	0.405 *** (0.146)	0.415 *** (0.153)	0.403 *** (0.140)
Colony	0.323 * (0.188)	0.402 * (0.214)	0.434 ** (0.214)	0.395 * (0.211)	0.397 * (0.217)	0.428 * (0.219)	0.352 (0.216)
Common language	0.437 *** (0.130)	0.546 *** (0.112)	0.552 *** (0.113)	0.538 *** (0.115)	0.572 *** (0.119)	0.555 *** (0.112)	0.560 *** (0.115)
Common colonizer	0.044 (0.191)	0.174 (0.303)	0.115 (0.293)	0.128 (0.311)	0.238 (0.309)	0.179 (0.312)	0.412 (0.309)
Same country	0.206 (0.285)	0.154 (0.229)	0.132 (0.233)	0.170 (0.222)	0.166 (0.223)	0.141 (0.235)	0.216 (0.232)
Intl	-1.912 *** (0.706)	-3.824 *** (0.789)	-4.833 *** (0.792)	-2.540 (1.560)	-6.136 *** (0.701)	-5.397 *** (0.313)	-2.824 *** (0.685)
Constant	8.256 *** (0.409)	7.311 *** (0.349)	7.355 *** (0.353)	7.308 *** (0.351)	7.362 *** (0.346)	7.329 *** (0.343)	7.516 *** (0.340)
Observations	8418	12444	12444	12444	12444	12444	12261
R2	0.958	0.952	0.952	0.952	0.952	0.952	0.953
Importer fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exporter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: All models are estimated by PPML with importer and exporter fixed effects. Robust standard errors adjusted for clustering by country pair. Statistical significance: * (10%), ** (5%), and *** (1%).

To convert these results into more usable form, we reformulate the trade costs function to express the elastic net SPI in ad valorem equivalent (AVE) terms, following Benz (2017):

$$AVE_j \equiv t_{ij} - 1 = \exp\left(\frac{-\beta SPI_j}{\theta}\right) - 1$$

Full results are presented in Appendix 2 (see on-line supplementary material). The advantage of presenting results in AVE terms is that it provides a simple number that can be compared across countries. It also shows that on a comparable standard, levels of restriction in services trade are typically higher than in goods. However, the AVE concept makes some significant

Table 10. Gravity models for total trade (goods and services), STRI and SPIs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	STRI	Elastic Net	Lasso	Ridge	PF	Mean	WB
STRI*Intl	−2.939 *** (0.649)	−2.475 *** (0.925)	−1.898 (1.182)	−4.445 *** (1.535)	2.385 ** (1.155)	−0.526 * (0.297)	−2.128 *** (0.798)
STRI*Intl*EIA	1.188 *** (0.347)	2.127 *** (0.400)	2.196 *** (0.420)	2.214 *** (0.440)	−1.085 (1.229)	0.508 *** (0.100)	0.015 *** (0.003)
Log(Tariff)	−0.283 (1.519)	−7.242 *** (1.833)	−7.942 *** (1.832)	−6.170 *** (1.870)	−12.551 *** (2.293)	−8.984 *** (1.954)	−6.897 *** (1.795)
PTA	0.074 (0.122)	−0.277 ** (0.119)	−0.305 ** (0.122)	−0.272 ** (0.122)	0.120 (0.117)	−0.320 *** (0.124)	−0.270 ** (0.120)
Log(distance)	−0.548 *** (0.059)	−0.443 *** (0.059)	−0.443 *** (0.059)	−0.442 *** (0.059)	−0.450 *** (0.059)	−0.439 *** (0.059)	−0.447 *** (0.059)
Contiguous	0.443 *** (0.126)	0.502 *** (0.138)	0.499 *** (0.139)	0.506 *** (0.135)	0.473 *** (0.162)	0.484 *** (0.145)	0.503 *** (0.135)
Colony	0.176 (0.148)	0.191 (0.138)	0.204 (0.137)	0.174 (0.136)	0.212 (0.134)	0.234 * (0.137)	0.171 (0.135)
Common language	0.159 (0.105)	0.322 *** (0.098)	0.327 *** (0.097)	0.315 *** (0.098)	0.316 *** (0.105)	0.336 *** (0.100)	0.318 *** (0.098)
Common colonizer	0.172 (0.138)	0.137 (0.329)	0.106 (0.332)	0.121 (0.331)	−0.092 (0.400)	0.050 (0.342)	0.111 (0.338)
Same country	0.609 *** (0.234)	0.744 *** (0.233)	0.737 *** (0.235)	0.762 *** (0.234)	0.783 *** (0.286)	0.734 *** (0.250)	0.781 *** (0.228)
Intl	−3.358 *** (0.248)	−3.565 *** (0.261)	−3.709 *** (0.312)	−3.052 *** (0.374)	−4.079 *** (0.199)	−3.591 *** (0.366)	−3.418 *** (0.361)
Constant	12.188 *** (0.369)	11.341 *** (0.366)	11.344 *** (0.364)	11.339 *** (0.364)	11.387 *** (0.362)	11.320 *** (0.366)	11.378 *** (0.364)
Observations	8366	12392	12392	12392	12392	12392	12209.000
R2	0.988	0.985	0.985	0.985	0.985	0.985	0.985
Importer fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exporter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Robust s.e. adjusted for clustering by country pair. Statistical significance: * (10%), ** (5%), *** (1%).

simplifications, including by assuming that the bundle of regulatory measures captured by the SPI has economic effects like those of a tariff. While that may be true for some measures, it may not be for others, especially those that affect competitive conditions in the marketplace or influence the fixed costs of market entry. We therefore present AVEs subject to these caveats.

While any indicator of services trade restrictiveness should be a strong predictor of bilateral services trade, recent work has shown that because of the input–output relationships that exist between services and other sectors, it is also likely that services policies affect total trade (i.e.,

goods and services).¹¹ We test this hypothesis and the predictive power of our SPIs compared with the STRI using aggregate Eora data summed across all 26 goods and services sectors in the database. The specification is the same as in the preceding tables, except that we use a dummy for PTA rather than EIA membership, to capture goods agreements as well as services agreements, and we include the log of the applied tariff rate as an additional explanatory variable. We aggregate the OECD STRI and our SPIs by taking simple averages across sectors, while the World Bank STRI is provided with a separate aggregate indicator.

Table 10 reports the results. We again use the full sample, but as tariff data are not available for all country pairs, sample size falls. Both STRIs, elastic net and ridge SPIs have the expected negative coefficients, and all four variables also have positive coefficients on the interaction term with the EIA variable, with all estimates statistically significant at the 1% level. The simple mean SPI also displays this pattern of coefficients, but the principal factor SPI has unexpected signs.

The World Bank STRI performs similarly to the OECD index or our SPIs. The only result that is contrary to expectations is the negative sign on PTA, but this is likely due to the fact that there is also an interaction term with EIA – an element of the PTA variable – that creates a correlation between the two. A similar issue arises in relation to tariffs, which are negatively correlated with PTA membership. Other variables perform largely in accordance with expectations.

The conclusion we draw in this section is that the OECD STRI, and the extended version with 23 additional countries that we have termed the SPI, is a robust predictor of bilateral trade at the sectoral level for services, and also at the aggregate level for goods and services. We interpret this as showing that the indices capture important features of services trade policy.

5. Conclusion

This paper provides new evidence on the state of services policies in 23 non-OECD countries in 2016, based on regulatory data released by the World Bank and WTO. Starting from the premise that the OECD STRI represents a proven approach to summarizing the restrictiveness of services policies, we use simple machine learning techniques to estimate SPIs for the new data that correlate very closely with OECD measures within sample, and therefore essentially constitute an extension of the OECD methodology to an additional set of mostly developing countries. Our SPIs provide the first quantitative snapshot of applied services policies in a significant number of developing countries since the World Bank's 2008–11 STRI.¹² They have significant explanatory power for bilateral trade flows at the sectoral and aggregate levels.

A contribution of this paper to the literature is to provide a proof of concept for the use of statistical tools, such as machine learning, to capture the complexities, nonlinearities, and dependencies of different services policy measures. This is relevant for at least two reasons. One is that the use of such techniques allows analysts to extend datasets in instances where a given source of information is limited to a subset of countries and the detailed methodology used to calculate published indicators is confidential. This is the case for the OECD STRI, arguably the current 'best in class' at the time of writing given extensive industry consultation and expert input into the weighting of measures across sectors. Insofar as other organizations – in this case the World Bank and WTO – collect similar types of policy data, SPIs that correlate well with the OECD STRIs offer a way to extend the country coverage of services restrictiveness indicators. Although our focus is on services trade restrictions, the methodology may be useful in

¹¹Hoekman and Shepherd (2017) and Shepherd (2019).

¹²Following the first version of the present paper, World Bank STRIs based on 2016 were reported in Borchert et al. (2020). These STRIs correlate closely with both the OECD index and our extension of it. Borchert et al. (2020) note that a comparison with 2008–11 STRIs reveals that many countries appear to have reduced the restrictiveness of applied policies somewhat. To facilitate more detailed analysis of the dynamics of policy change around the world, it would be desirable for the original policy data used to generate the World Bank 2008–11 STRIs to be made available in comparable format through the I-TIP platform.

other contexts where similar conditions prevail as regards the scope and periodicity of efforts to collect information on policies for a given area.

More generally, the use of statistical tools may help to identify potential ways to reduce data collection costs. The OECD STRI involves the collection of a large amount of data, entailing significant direct and time costs for agencies. Further work with machine learning algorithms like those deployed here may identify a subset of measures that in fact do most of the explanatory work. In our view, this is the primary value of generating these kinds of indices, rather than simply summarizing a vast amount of data in a single number. Data collection is distinct from research to fine-tune STRI methodologies and improve the associated weighting and aggregation measures. The latter is very important but should be independent of the policy collection process. Analysts should have the ability to define their own indicators, and it is therefore very welcome that I-TIP has released the 2016 services policy information independently of associated STRIs.

Although the release of services trade-related policy data in I-TIP is laudatory, as of 2020 the most up-to-date compilation of such measures will be for 2016, and then only for some 30 developing countries – without any coverage of most low-income countries. It is unknown whether and when a new wave of data will be collected and thus whether over time a panel dataset will emerge. Greater effort to generate services policy data on a regular basis for a broad range of countries would allow governments to track their policies, compare them to those of other countries, and inform autonomous policy reforms and regional integration processes. The resource costs of a systematic effort to collect services policy data are not large. Focusing on just five major sectors per country and seeking to cover 50 non-OECD countries would, based on our prior experience (e.g., Shepherd et al., 2019a,b), involve costs in the range of \$400,000 to \$750,000, with additional resources required for reporting and publishing, though they would be an order of magnitude less than those required for data compilation. Doubling coverage to 10 sectors would involve an investment of less than \$2 million. Average costs could be reduced by making the data collection a bi-annual process. Similarly, application of machine learning to reduce the burden of data collection could perhaps reduce these numbers by half.

Given how limited services policy data are relative to information on merchandise trade policies, allocating this level of resources to filling the gap would have a very high cost–benefit ratio, especially if one considers the opportunity costs of not having up-to-date information on services policies. These opportunity costs may be high, not least because absence of data means policy-makers may be less inclined to devote adequate attention to this important area of policy.¹³ Similarly, important research questions, such as the possibility for heterogeneous policy effects across different countries, cannot be investigated in the absence of additional data, collected and disseminated at regular intervals.

One priority in this regard is to incorporate the preferential dimension into measures of services policy restrictiveness. Another is to expand country coverage. In particular, very few African countries are included in I-TIP. Given the salience of regional integration in Africa, it is important to fill in the policy blanks to allow assessments of the utility of dealing with services in the context of pursuing continental free trade. Benz and Gonzalez (2019) have shown that the EU single market for services is much more liberal than any member country's MFN policies. The extent to which other trade agreements effectively liberalize services markets is unclear but is a vital policy question in an environment where bilateral, plurilateral, and mega-regional agreements are becoming more common. Shepherd et al. (2019a) find little evidence of substantial liberalization in the Canada–EU Trade Agreement (CETA). The same appears to be true for the Comprehensive and Progressive Trans-Pacific Partnership (CPTPP) (Gootiiz and Mattoo, 2017).

A related important question concerns the value of making binding policy commitments in trade agreements, even if these do not entail liberalization. The 'water' in the services policy

¹³Other compilations of policy indicators such as the World Bank Doing Business project attract extensive attention by the press and have become focal points for governments because they are undertaken on an annual basis.

commitments in trade agreements often is considerable (see, e.g., Borchert, Gootiiz, and Mattoo, 2011; Miroudot and Shepherd, 2014; Miroudot and Pertel, 2015; Ciuriak et al., 2017). Research on the value of reducing the difference between bound and applied services policies has shown that this may be an important source of welfare gain, driven by a reduction in policy uncertainty (Lamprecht and Miroudot, 2018; Ciuriak et al., 2019; Egger et al., 2019).

Again, such analysis requires good quality, comparable information on applied policies collected regularly. The OECD does this for its members – and is the source for the majority of the 68 countries for which I-TIP reports comprehensive information. Looking forward, we hope this gap will be filled for developing countries by agencies with appropriate mandates.

Supplementary Materials. To view supplementary material for this article, please visit <https://doi.org/10.1017/S1474745620000439>.

Acknowledgements. We are grateful to the Editor (Alan Winters) and two reviewers for helpful comments and constructive suggestions on an earlier draft.

References

- Anderson, J. and E. Van Wincoop (2003) 'Gravity with Gravitas: A Solution to the Border Puzzle', *American Economic Review* 93(1), 170–192.
- Anderson, J. and E. Van Wincoop (2004) 'Trade Costs', *Journal of Economic Literature* 42(3): 691–751.
- Anderson, J., M. Larch, and Y. Yoto (2018) 'GPEPML: General Equilibrium Analysis with PPML', *The World Economy* 41 (10), 2750–2782.
- Arkolakis, C., A. Costinot, and A. Rodriguez-Clare (2012) 'New Trade Models, Same Old Gains?', *American Economic Review* 102(1), 94–130.
- Athey, S., M. Bayati, N. Doudchenko, G. Imbens, and K. Khosravi (2017) 'Matrix Completion Methods for Causal Panel Data Models', Working Paper, arXiv: 1710.10251, <https://arxiv.org/abs/1710.10251>.
- Athey, S. and G. Imbens (2019) 'Machine Learning Methods that Economists Should Know', *Annual Review of Economics* 11, 685–725.
- Benz, S. (2017) 'Services Trade Costs: Tariff Equivalents of Services Trade Restrictions Using Gravity Estimation', OECD Trade Policy Paper 200.
- Benz, S. and F. Gonzales (2019) 'Intra-EEA STRI Database: Methodology and Results', OECD Trade Policy Paper 223.
- Beverelli, C., M. Fiorini, and B. Hoekman (2017) 'Services Trade Policy and Manufacturing Productivity: The Role of Institutions', *Journal of International Economics* 104, 166–182.
- Borchert, I., B. Gootiiz, and A. Mattoo (2011) 'Services in Doha: What's on the Table?', in W. Martin and A. Mattoo (eds.), *Unfinished Business: The WTO's Doha Agenda*. Washington, DC: World Bank & CEPR.
- Borchert, I., B. Gootiiz, and A. Mattoo (2014) 'Policy Barriers to International Trade in Services: Evidence from a New Dataset', *World Bank Economic Review* 28(1), 162–188.
- Borchert, I., J. Magdeleine, J. Marchetti, and A. Mattoo (2020) 'The Evolution of Services Trade Policy since the Great Recession', *WTO Staff Working Paper ERSD-2020-02*.
- Borchert, I., B. Gootiiz, J. Magdeleine, J. Marchetti, and A. Mattoo (2019) 'Restrictions on International Trade in Services: Evidence from a New Database', GTAP Working Paper, www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=5537.
- Chaney, T. (2008) 'Distorted Gravity: The Intensive and Extensive Margins of International Trade', *American Economic Review* 98(4), 1707–1721.
- Ciuriak, D., J. Xiao, and A. Dadkhah (2017) 'Quantifying the Comprehensive and Progressive Agreement for Trans-Pacific Partnership', *East Asian Economic Review* 21(4), 343–386.
- Ciuriak, D., A. Dadkhah, and D. Lysenko (2019) 'The Effect of Binding Commitments on Services Trade', *World Trade Review*, doi:10.1017/S1474745618000496.
- Costinot, A. and A. Rodriguez-Clare (2014) 'Trade Theory with Numbers: Quantifying the Consequences of Globalization', in G. Gopinath, E. Helpman, and K. Rogoff (eds.), *Handbook of International Economics*, Vol. IV. Amsterdam: Elsevier.
- Eaton, J. and S. Kortum (2002) 'Technology, Geography, and Trade', *Econometrica* 70(5), 1741–1779.
- Egger, P. and M. Larch (2008) 'Interdependent Preferential Trade Agreement Memberships: An Empirical Analysis', *Journal of International Economics* 76(2), 384–399.
- Egger, P., J. Francois, B. Hoekman, and M. Manchin (2019) 'Regulatory Bindings, Policy Uncertainty and Market Access in Services', in J. Francois and B. Hoekman (eds.), *Behind-the-Border Policies: Assessing and Addressing Non-Tariff Measures*. Cambridge University Press.
- Fally, T. (2015) 'Structural Gravity and Fixed Effects', *Journal of International Economics* 97, 76–85.

- Fiorini, M. and B. Hoekman (2018) 'Services Trade Policy and Sustainable Development', *World Development* 112 (December), 1–12.
- Francois, J. and B. Hoekman (2010) 'Services Trade and Policy', *Journal of Economic Literature* 48(3), 642–692.
- Gootiz, B. and A. Mattoo (2017) 'Services in the Trans-Pacific Partnership', World Bank Policy Research Working Paper 7964.
- Grosso, M., F. Gonzales, S. Miroudot, H. Nordås, D. Rouzet, and A. Ueno (2015) 'Services Trade Restrictiveness Index (STRI): Scoring and Weighting Methodology', OECD Trade Policy Paper 177.
- Heid, B., M. Larch, and Y. Yotov (2017) 'Estimating the Effects of Non-Discriminatory Trade Policies within Structural Gravity Models', Working Paper 2017-10, Drexel University LeBow College of Business.
- Helble, M. and B. Shepherd (eds.) (2019) *Leveraging Services for Development: Prospects and Policies*. Tokyo: ADBI.
- Hoekman, B. and B. Shepherd (2017) 'Services Productivity, Trade Policy and Manufacturing Exports', *The World Economy* 40(3), 499–516.
- Jafari, Y. and D. Tarr (2017) 'Estimates of Ad Valorem Equivalents of Barriers against Foreign Suppliers of Services in Eleven Services Sectors and 103 Countries', *The World Economy* 40(3), 544–573.
- Lamprecht, P. and S. Miroudot (2018) 'The Value of Market Access and National Treatment Commitments in Services Trade Agreements', OECD Trade Policy Papers 213.
- Miroudot, S. and K. Pertel (2015) 'Water in the GATS: Methodology and Results', OECD Trade Policy Paper 185.
- Miroudot, S. and B. Shepherd (2014) 'The Paradox of Preferences: Regional Trade Agreements and Trade Costs in Services', *The World Economy* 37(12), 1751–1772.
- Nordås, H. (2016) 'Services Trade Restrictiveness Index (STRI): The Trade Effect of Regulatory Differences', OECD Trade Policy Paper 189.
- Nordås, H. (2018) 'What Drives Trade in Services? Lessons from the Nordics', *Applied Economics* 50(33), 3532–3545.
- Nordås, H. and D. Rouzet (2017) 'The Impact of Services Trade Restrictiveness on Trade Flows', *The World Economy* 40(6), 1155–1183.
- OECD (2019) 'Services Trade Restrictiveness Index', www.oecd.org/tad/services-trade/services-trade-restrictiveness-index.htm.
- Shepherd, B. (2019) 'Services Policies and Manufacturing Exports', in M. Helble and B. Shepherd (eds.), *Leveraging Services for Development: Prospects and Policies*. Tokyo: ADBI.
- Shepherd, B., M. Décosterd, C. Castillo Comabella, and D. Stivas (2019a) 'EU Exit and Impacts on Northern Ireland's Services Trade', Report prepared for the Department for the Economy, Northern Ireland, www.economy-ni.gov.uk/sites/default/files/publications/economy/STRI-research-EU-Exit-and-Impact-on-NI-Services-trade.pdf.
- Shepherd, B., M. Décosterd, A. Markitanova, H. Mukama, A. Nderitu, and D. Stivas (2019b) *Quantifying the Barriers to Services Trade in the Commonwealth: A Focus on Kenya and Rwanda*. London: Commonwealth Secretariat.
- Su, X., S. Anwar, Y. Zhou, and X. Tang (2019) 'Services Trade Restrictiveness and Manufacturing Export Sophistication', *North American Journal of Economics and Finance*, doi.org/10.1016/j.najef.2019.101058.

Appendix 1: I-TIP Additional Country Coverage

East Asia & Pacific	Europe & Central Asia	Latin America & Caribbean	Middle East & North Africa	South Asia	Sub-Saharan Africa
Hong Kong SAR, China	Kazakhstan	Argentina	Egypt	Bangladesh	Kenya*
Myanmar	Ukraine	Dominican Republic	Oman	Pakistan	Nigeria
Philippines		Ecuador	Tunisia	Sri Lanka	Rwanda*
Singapore		Panama			
Taiwan, China		Peru			
Thailand		Uruguay			
Vietnam					

Note: The table includes only those countries covered by the SPIs that are not included in the OECD STRI (OECD, 2019).

*Rwanda is not in I-TIP but comparable policy data for Rwanda were collected by Shepherd et al. (2019b), permitting its inclusion in the analysis. Kenya is included in I-TIP but data have been augmented by additional information reported in Shepherd et al. (2019b).

Cite this article: Hoekman B, Shepherd B (2020). Services Trade Policies and Economic Integration: New Evidence for Developing Countries. *World Trade Review* 1–20. <https://doi.org/10.1017/S1474745620000439>