ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics

Master Thesis International Economics

Trade and the Economic Complexity of Nations

Name student: Jose Mari Romuald Mayuga Tuaño
Student ID number: 380692

Supervisor: dr. Aksel Erbahar
Second assessor: prof. dr. Maarten E. Bosker

Date final version: 10/08/2017

Abstract

This paper studies the relationship between trade and the economic complexity of nations. Using data spanning between 1996 and 2015 on the economic complexity index (Simoes & Hidalgo, 2011), various macroeconomic controls, and governance indicators, this research uses a combination of static econometric approaches as well as the Arellano-Bond dynamic panel estimator. Although, there is some evidence for a relationship between economic complexity and the total value of imports and exports, the relationships themselves are likely to be defined by a variety of opposing and reinforcing effects. In contrast, there appears to be a rather weak relationship between weighted-mean applied tariff rates and economic complexity. Nonetheless, this paper simply scratches the surface of the matter and future efforts in studying this relationship is likely to face a number of obstacles.

I – Introduction

Let us begin with a visualisation of David Ricardo's (1891) historical example of international trade. Here, differences in relative labour productivity lead to England's specialisation in the production of cloth and Portugal's specialisation in wine. In the framework of the Ricardian model, countries benefit from the propagation of trade because specialisation allows the world to benefit from the production of more goods as a whole. However, do the products that countries export specifically matter for their future development?

Hidalgo et al. (2007) find that by modelling similarities across different products – via the "product space", they can show that a country's exports *can* condition future economic development. England's specialisation in cloth implies that its future industrial development lies in products that share similarities in the production process of cloth. For example, other types of textiles – such as articles of clothing or industrial cables and ropes. On the other hand, Portugal's specialisation in wine may lead to fewer opportunities for future industrial development. This is because the production process of wine has fewer industrial applications other than in the production of other foodstuffs. Further developing this idea, Hidalgo and Hausmann (2009) introduce the Economic Complexity Index (ECI). The ECI takes a holistic approach to international trade and specialisation, highlighting the fact that wealthier and more developed economies tend to have a highly diversified portfolio of exported goods which are exported by few other countries. Returning to the example of England and Portugal, since Portugal may have had fewer future development opportunities, England conversely may have had an opportunity to become relatively wealthier in the future.

Therefore, one may ask, how does a country's openness to trade or trading activities affect its degree of economic complexity? As discussed by Dany Bahar (2016), "While economic theory suggests that specialisation is a result of openness to trade, it is less clear what is the general equilibrium outcome of integration in the global economy". Here, the general equilibrium outcome of integration refers to the degree of economic complexity. More specifically, does increasing trade activity or liberalisation broaden the available opportunities for economic development? Or do patterns of specialisation in accordance with standard trade theory actually limit those opportunities?

Thus, the research question of this paper can be broadly stated as: "What is the relationship between trade and economic complexity?"

It is important to investigate this relationship due to the implications associated with ECI as a macroeconomic performance measure. For example, the ECI can provide insight into social outcomes such as income inequality (Freitas and Paiva, 2016; Hartmann et al., 2017a; Hartmann et al., 2017b). In many ways, the holistic approach of ECI and its underlying concepts provide additional detail which traditional macroeconomic variables may be unable to provide. For instance, GDP can only measure the total "value" of what a country exports, whereas ECI can account for "what" those exports are and what this may mean for its future. For example, Dutch disease describes a situation in which the economic development of a valuable natural resource sector leads to a decline in other sectors. Here, GDP trends may indicate a fruitful economic situation but ECI may not. A good example is the case of Venezuela – where the GDP per capita and ECI trends preceding its current socio-economic crisis[1] indicate starkly different implications.

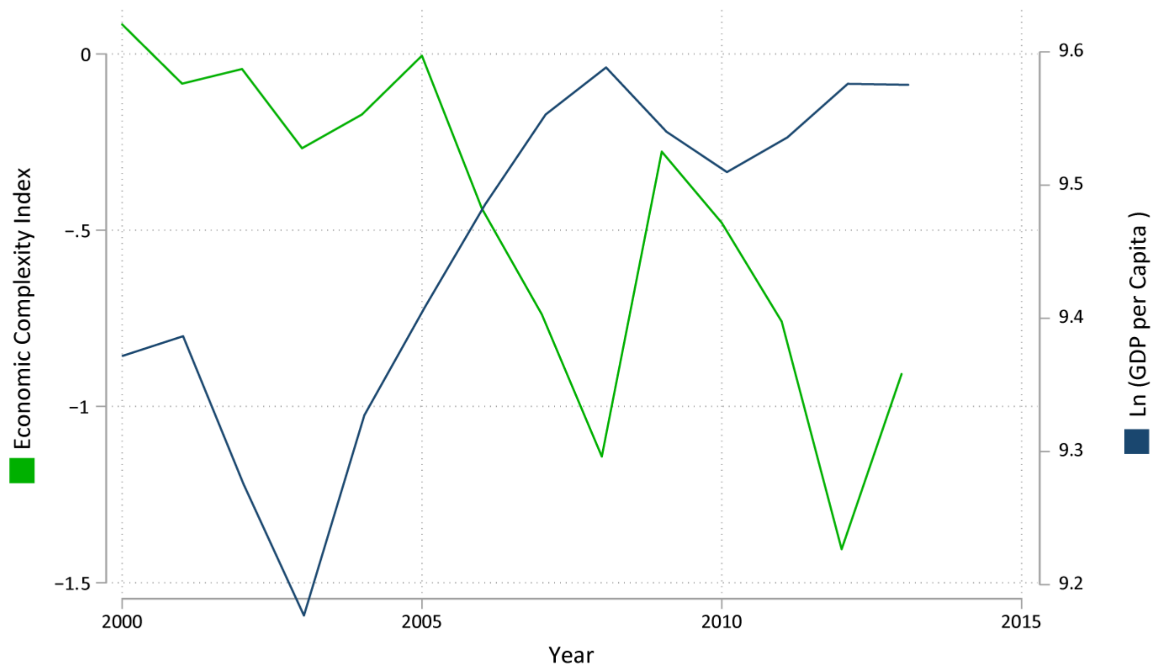


Figure 1: ECI and Ln (GDP per capita) by Year for Venezuela

Note that although the crisis began in 2012, the overall trend for the economic complexity of Venezuela indicates that issues regarding their productive structure began since the year 2000. This is a fact that is not captured by the natural log of GDP per capita, where the overall trend over the same period appears to be increasing. Though this is an oversimplified

[1] https://www.economist.com/blogs/americasview/2012/08/venezuelas-oil-industry?zid=305&ah=417bd5664dc76da5d98af4f7a640fd8a

example, the stark difference in trends clearly shows the disparity between both measures as well as the importance of detail in macroeconomic analysis.

The remainder of this paper is structured as follows. Section II discusses the related literature, Section III describes the data, outlines the methodology to be used, and conducts some preliminary analysis, whereas Section IV presents the main results. Section V reflects upon the results, discussing its implications and limitations. Section VI concludes the paper and provides some suggestions for future research.

II – Related Literature

II. A. Pioneers of Productive Structure and Development Economics

The idea that the type of products that a country exports matters for subsequent economic performance is not new. The "fathers" of development economics were the first to discuss this, emphasising the role of productive structure and its diversification in economic growth (Rosenstein-Rodan, 1943; Hirschman, 1958). Paul Rosenstein-Rodan (1943), for example, states that the complementarities between different industries are important in planning large-scale industrialisation. Although these early concepts of productive structure share some common ground with the more recent literature measuring economic complexity, research on productive structures was severely limited due to difficulties in empiric measurement. At the time, economic data lacked the disaggregated detail that is required to effectively quantify and define the productive structure of a country.

Indeed, the most significant early economic works in terms of measuring productive structure are limited to aggregated measures and concepts. For example, the division of a country's industries into primary, secondary, or tertiary sectors. Furthermore, measures of aggregate concentration and diversification such as the Herfindahl-Hirschman Index (Hirschman, 1945; Herfindahl, 1950) face difficulties in capturing the complexity of products and thus, overlook the detail of industrial structures. Other related research distinguishes between products of a related and unrelated variety, finding evidence of evolutionary notions that economic development and international trade patterns are path dependent (Frenken, van Oort, & Verburg, 2007; Saviotti & Frenken, 2008; Boschma & Iammarino, 2009).

II.B. Capability Theory and Economic Complexity

Recently, a theory of capabilities for productive transformation was published by the International Labour Organization (ILO) (Nübler, 2014) in an attempt to formally combine the related literature under a coherent theory[2]. Nübler (2014)'s framework is extensive and thus, it is useful to focus on its most relevant elements. Nübler emphasises the importance of productive capabilities, which are considered embodiments of various collective forms of knowledge. This definition separates them from productive capacities which instead reside in the "material" sphere of an economy, i.e. tangible production factors. Thus, the process of acquiring capabilities is a process of learning. Nübler's capabilities framework, therefore, captures the importance of learning, routines and institutions in productive transformation and economic development. For example, having an abundance of iron in an economy does not necessarily imply that the economy has the capability to produce steel – which requires knowledge in both the process of smelting iron as well as alloying.

The economic complexity index (ECI) introduced by Hidalgo and Hausmann (2009) focuses on exports-based measurement of a country's productive structure. It has allowed for the quantification of the complexity of a nation's productive structure, further contributing to the literature on the patterns of structural transformations. This specific strand of literature begins with the idea that what a country produces and consequently exports, matters for its economic development (Hausmann, & Klinger, 2006; Hausmann, Hwang, & Rodrik, 2007; Rodrik, 2006; Hidalgo et al., 2007; Hidalgo, 2009; Hausmann & Hidalgo, 2011; Bustos et al., 2012). Hidalgo et al. (2007) were the first to establish the formal framework in which the "product space" lies. Here, products are assumed to be related if they require similar capabilities, and thus will tend to be produced and exported in tandem. The key observation in this product space is that a country cannot simply produce whatever they would like. For example, one cannot simply "move" from exporting crude oil to exporting apples even when they have the resources to do so because the capabilities required to produce apples are completely different. Instead, it would be easier to shift towards exporting refined oil, which shares many more required capabilities with crude oil. Therefore, the resulting product space "maps" out the relationships between industries, and can be used to observe the status and evolution of a country's productive structure.

Building on these findings, Hidalgo and Hausmann (2009) establishes the formal basis for their measure of economic complexity. The main connection between ECI and the product space is that the rigidity of the product space is what prevents countries from improving its economic complexity from year to year. The ECI is measured by inferring information from

[2] Note that this "productionist view" of capabilities is different from the "humanistic view" developed by Amartya Sen.

the matrix of products exported by a country with RCA[3] - which can be used to determine the *diversification* of an economy's exports, as well as the *ubiquity* of those products. For example, a highly diversified country is said to produce and export a large basket of unique goods and a highly ubiquitous good is said to be produced and exported by a large set of unique countries. Using these metrics, the authors are able to develop an interpretation of the network relationship between countries, products, and capabilities. In relation to the framework laid out by Nübler (2014), Hidalgo and Hausmann (2009, p. 1) views the bipartite trade network between countries as "…the result of a larger, tripartite network, connecting countries to the capabilities they have and products to the capabilities they require." The authors are able to show that economic complexity captures information about a country's productive structure, strongly correlates with income per capita, and is predictive of future economic growth[4]. Furthermore, the authors consider that a country's future economic development is intrinsically linked to its economic complexity due to the strong path dependency exhibited by countries' productive structures.

This approach of economic complexity spurred on research in several other areas. Economic complexity's characteristic of path dependence highlights its usefulness as a tool to identify opportunities for structural transformation in developing countries and thus future economic growth (Hidalgo, 2011; Abdon & Felipe, 2011; Felipe, 2012; Hausmann et al., 2014). Other publications build on the methods used by Hidalgo and Hausmann (2009), making their own improvements in the measurement of economic complexity (Caldarelli et al., 2012; Cristelli et al., 2013; Cristelli, Tacchella, & Pietronero, 2015). Finally, the most recent literature on economic complexity expands its scope into areas of income inequality (Freitas & Paiva, 2016; Hartmann et al., 2017a; Hartmann et al., 2017b).

II.C. Trade Literature

Since this paper is the first (as is to my knowledge) to study the relationship between trade and economic complexity, there are no comparable studies to set a benchmark for results. However, the literature on trade and growth is possibly one of the most studied and scrutinised

[3] Revealed comparative advantage (RCA) refers to the Balassa Index, introduced by Balassa (1965). It measures the "relative advantage" of a country in a certain classification of goods using trade flow data.
[4] Hidalgo and Hausman (2009, p. 5) do this by regressing the rate of growth of income per capita on different measures of economic complexity and on a country's initial level of income. Results are valid for a 20-year period (1985-2005), two 10-year periods, and four 5-year periods.

areas in international economics. It would, therefore, be useful to pay attention to the existing trade literature and how their findings may shed some initial insights into this study.

Possibly some of the most famous and perhaps, infamous, studies that link trade liberalisation with economic growth are the studies conducted by Dollar (1992), Sachs and Warner (1995), Edwards (1998), and Frankel and Romer (1999). Although all four cross-sectional studies find strong evidence for the presence of a positive relationship between trade liberalisation/openness with economic growth, they have also received strong criticism. The main issues pointed out by other researchers primarily lie in the choice of measures of openness or liberalisation, as well poor econometric methodology (Rodriguez and Rodrik, 2000). Winters (2004) lays out four main issues when studying the empirics between liberal trade and growth. The first issue is that the choice of trade liberalisation proxies is subjective. For instance, where Dollar (1992) relies on the use of the volatility of the real exchange rate, Sachs and Warner (1995) instead create a composite measure combining tariffs, non-tariff measures, black market exchange rate premia, socialist tendencies, and the monopolisation of exports. Secondly, and similar to the first, comparing trade policy stances across countries is quite difficult. Thirdly, causation is extremely difficult to establish. Frankel and Romer (1999) were one of the first to attempt establishing a causal relationship by instrumenting their measure of openness with geographic factors. Even then, however, Rodriguez and Rodrik (2000) postulate that even geographical variables could have effects on growth. For example, geography may well have effects on endowments or even institutions. Finally, for liberal trade policies to have a long-term effect on growth, they are likely to require being combined with other good policies, for instance, investment policies. The primary problem here is that it is difficult to accurately disentangle the relative importance of other macroeconomic policies in relation to those of liberal trade.

Assessing the tripartite relationships between trade, growth and other factors can also be of interest when determining which variables to consider as covariates. Taylor (1998) and Wacziarg (2001), for instance, identify good investment policy as a key link between liberal trade policy and positive outcomes on economic growth. Furthermore, institutional factors are considered to have a considerable effect on GDP growth (Acemoglu, Johnson, & Robinson, 2000; Rodrik, Subramanian, & Trebbi, 2004; Dollar & Kraay, 2003). Rodrik, Subramanian, and Trebbi (2004) argue that institutions outperform trade openness as an explanation of real income per head and that given institutions, openness has an insignificantly negative effect. In the context of this study, it is not difficult to consider that these factors – investments and institutions – may also have an effect on the economic complexity of a nation.

Returning to the issues outlined by Winters (2004), this paper is likely to share many similar obstacles with respect to establishing a relationship between economic complexity and trade. The first challenge is to select this paper's measure of trade openness or liberalisation. As proxies, the total values of imports and exports and weighted-average applied tariffs are intuitive choices at the aggregate country-level. Since the total values of imports and exports directly measure the value of trade that a country engages in, this set of variables can be interpreted as a measure of "direct" trade openness. However, its use in this study requires the assumption that increases in trade activity are a direct consequence of trade liberalisation policy. Clearly, such an assumption is fragile given the broad spectrum of factors that are likely to determine trade activity. The alternative approach, proxying directly for trade policy, can use weighted-average applied tariff rates. Since tariffs are a protectionist policy, the level of tariffs provides an indication of how liberal a country's stance is towards trade. Moreover, since countries may be members of a wide set of different trade agreements or trade associations, tariff measures can be distinguished as "Most-Favoured-Nation" (MFN) or "Effectively Applied" tariffs. MFN tariffs are the rates charged to imports from other members of the WTO. However, if two sets of countries have signed a preferential trade agreement, the applied tariff can differ. Therefore, to simplify matters, the effectively applied tariff is defined as the lowest tariff applied by the importing nation. Furthermore, a weighted average applied tariff was chosen over a simple average in order to account for the imported product shares from each partner country. Conceptually, this should provide a more accurate assessment of trade liberalisation as it places less weight on unimportant trading partners.

However, one must also take into account the presence of non-tariff barriers (NTBs). NTBs range from other common protectionist tools such as import quotas to more immeasurable barriers such as local content requirements. In cases of instrument substitution, reductions in tariffs may coincide with increases in the use of NTBs. In general, it is a considerable challenge to coherently measure the coverage of NTBs in any given country given the diversity of NTB types and the ambiguity in weighing each type's importance. For example, although the WITS[5] database records data on NTBs, the data does not match the desired aggregate panel-structure of this study. More specifically, the data is disaggregated at the

[5] http://wits.worldbank.org/WITS/WITS/QuickQuery/NTMs/NTMsQuickQuery.aspx?Page=NTMs

product level and is also differentiated by NTB type. Combining this information in a coherent way, at an aggregate level, without any indication of the relative importance of NTB types is a significant challenge and would be beyond the scope of this paper.

Delving deeper into the efforts of researchers, there exists a wide variety of alternative strategies to proxy for trade liberalisation. In accordance with classical trade theory, trade liberalisation can be associated with convergence to relative price neutrality. The most prominent examples of such an approach include Krueger's (1978) bias measure and Balassa's (1982) relative exchange rate measure. However, these measures to be particularly data-intensive, limiting their usefulness in this study. One other approach involves identifying a particular event or action that signals a nation's intent to liberalise trade. For example, the signing of an important FTA or a country's first World Bank structural adjustment loan (SAL)[6]. However, intent does not always translate into real trade liberalisation. Finally, some researchers attempt to combine multiple criteria on a wide range of indicators. As mentioned earlier, Sachs and Warner (1995) take this approach, leading to the creation of a conceptually rich measure of liberalisation. However, as with any composite measure, the relative weights and importance of each criterion tend to be subject to the judgement of its creator(s).

In conclusion, there is no clear-cut winner between indicators. Choices are generally made on the basis of a researcher's judgement, and an indicator's feasibility and convenience. Therefore, to remain within the scope of this research, the choices of proxies here are based on convenience, data availability, and their ease of interpretation. The proxies of total trade value and weighted-average applied tariffs will be used separately in the specifications that follow in order to observe outcomes under different perspectives on trade. However, it is paramount that conclusions are drawn from these specifications be interpreted with care. First, changes in weighted-average applied tariffs can be roughly interpreted as changes in trade policy, whereas changes in the total value of imports or exports are interpretable as changes in direct trade openness. Secondly, neither measure chosen here is insusceptible to issues of endogeneity. For example, tariff rates may be related to political pressures associated with higher levels of income inequality or lower rates of economic growth – both of which were shown to relate to ECI (Hartmann et al., 2017a; Hidalgo & Hausmann, 2009).

[6] Greenaway et al. (2002) make use of a country's first World Bank SAL or equivalent intervention as one of three different indicators of liberalisation.

III – Data and Methodology

III. A. Data Descriptions and Sources

The sole dependent variable in this study is the Economic Complexity Index (ECI) (Simoes & Hidalgo, 2011). ECI data over the period 1964 – 2015 was acquired from the Observatory of Economic Complexity (OEC)[7]. Data on the independent variables used in this research was obtained from World Bank Databases. More specifically, the World Development Indicators (2017) (WDI)[8] and the Worldwide Governance Indicators (2016) (WGI)[9] databases. The WDI database contains a large array of data ranging from trade data to foreign direct investment (FDI) data over the period 1960 – 2015. The WGI database contains aggregate indicators for six different dimensions of governance over the period 1996 – 2015.

As discussed earlier in the literature review section, the Economic Complexity Index (ECI) (Simoes & Hidalgo, 2011) measures the complexity of a country's economy – i.e. its productive structure. As a reminder, Hidalgo and Hausmann (2009) define economically complex countries as having a diverse export basket, and when those export products are less ubiquitous. Therefore, under the context of this study, countries with large positive values of ECI are considered relatively more complex, whereas those with large negative values are relatively simple.

The total values of both imports and exports (in constant 2010 US$) and the weighted mean applied tariff rates for each country are used as the main independent variables of interest. Import and export values are used in tandem to represent direct trade openness, whereas the tariff data proxies for trade policy. FDI inflows and outflows are used as controls for the effect of investment patterns on the ECI – where inflows are considered to be more conceptually relevant. For example, FDI inflows can be intuitively associated with investments into a country's productive structure, whereas FDI outflows can be inversely thought of as a result of a country's productive structure. Other controls include GDP per capita (in constant 2010 US$), which has been identified by Hidalgo and Hausmann (2009) as strongly correlated to the level of ECI in a country, total population as a proxy for country size, and service sector size to control for the importance of services in a country's economy. Digressing shortly on the relevance of the service sector, its inclusion is largely motivated by limitations in the

[7] http://atlas.media.mit.edu/en/
[8] http://data.worldbank.org/data-catalog/world-development-indicators
[9] http://data.worldbank.org/data-catalog/worldwide-governance-indicators

calculation of ECI. According to the research of Hartmann et al. (2017a), the estimated economic complexity of Australia is likely to be underestimated due to the exclusion of traded services in its estimation. By controlling for service sector size, we may be able to uncover an important relationship with economic complexity that may apply beyond the case of Australia. Finally, the variables “Control of Corruption”, “Government Effectiveness”, “Political Stability, No Violence”, “Rule of Law”, “Regulatory Quality”, and “Voice and Accountability” all represent six distinct areas of governance. All six variables are used to broadly control for governance heterogeneity across countries, and range from values of -2.5 (weak) to 2.5 (strong) to indicate governance quality. More detailed descriptions of the data can be found in the appendix under Table A1.

III. B. Summary Statistics and Transformations

Table A2 in the appendix summarises all the variables used in this study, which forms an unbalanced panel data ranging from the years 1960 up until 2015. From table A2, it is easy to see that different specifications in the coming analysis will make use of a starkly different number of observations. For example, specifications based on tariff data will have less information to work with since World Bank tariff data is only available from 1988 onwards. Furthermore, tariff data tends to have more missing observations for most countries than data on trade flows. The use of governance indicators as controls similarly reduces sample size, given that observations are only recorded from 1996 onwards. In an attempt to alleviate the issues in comparing specifications with large differences in the number of observations, the following analysis limits the sample period between 1996 and 2015. This adjusted sample now spans 20 years across 82 countries[10] and begins after important world events related to trade. For example, the collapse of the Soviet Union, the formation of the European Economic Area (EEA) and the formation of the WTO. By limiting the sample period to this time span, we avoid comparing model specifications across different world eras while still retaining a relatively large number of observations. The list of countries included in the sample is displayed in table A3 in the appendix.

Table A2 also describes which variables were subject to a natural logarithmic (ln) transformation. Note that although the primary use of the ln transformation was to reduce the effect of outliers on estimates, the application of the transformation was not always conceptually permitted. For example, variables that are measured in percentages or as an index

[10] The list of countries included in this adjusted sample can be found under table A3 in the appendix.

were not transformed. The variables for FDI flows, however, are a peculiar case. Although it is clear that there is evidence of right-skewness when comparing the variables' mean and median, the presence of negative and zero values prevents the use of a logarithmic transformation. At first glance, it may seem contradictory to have negative values of inflows or outflows since both should be strictly positive by definition. However, negative net inflows or outflows are in fact instances of divestment and thus should not be omitted.

III. C. Methodology

III. C. i. Static approaches

First, basic ordinary least squares (OLS) estimates are used to set a simple benchmark for later results. The two specifications using trade values and weighted mean applied tariffs, respectively, can be written as follows:

$$ECI_{i,t} = \beta_0 + \beta_1 \ln(\text{IM})_{i,t} + \beta_2 \ln(\text{EX})_{i,t} + \sum_{j=3}^{13} \beta_j X_{i,t} + u_{i,t} \quad \text{(E.1)}$$

$$ECI_{i,t} = \beta_0 + \beta_1 \text{TAWM}_{i,t} + \sum_{j=2}^{12} \beta_j X_{i,t} + u_{i,t} \quad \text{(E.2)}$$

Where the subscript "i" represents an individual country, "t" represents the year of the observation, and "$u_{i,t}$" is the error term. "ECI" is the shorthand for the economic complexity index, "ln (IM)" and "ln (EX)" are the shorthand representations of the natural logs of a country's total value of imports and exports, and TAWM is the shorthand for "Tariff rates, Applied, Weighted Mean". The set of control variables, "$X_{i,t}$", includes the natural logarithms of GDP per capita, total population, a country's service sector size, FDI inflows and outflows, and all six governance controls retrieved from the WGI database.

The second static approach makes use of a within Fixed Effects (FE) estimator. A commonly used method in a panel data context, the inclusion of within fixed effects allows the model to take account of independent effects for each country. It may be relevant to note that, by assumption, the FE model allows the fixed effects to be correlated with the independent regressors. A random effects (RE) estimator instead assumes that the random effects are uncorrelated with the other covariates. Later applied Hausman tests, however, indicate that the FE estimator is the more appropriate estimation approach (Table A9). In general, the FE estimator improves on OLS estimates by controlling for differences across countries by eliminating the fixed effects through a demeaning process. More specifically, each variable observation for a certain country in a given year is subtracted by the mean of that variable for the entire sample period [E.g. $\ddot{ECI}_{i,t} = ECI_{i,t} - \overline{ECI_i}$]. The resulting specifications are as follows:

$$\ddot{ECI}_{i,t} = \ddot{\beta}_0 + \beta_1 \ln(\ddot{IM})_{i,t} + \beta_2 \ln(\ddot{EX})_{i,t} + \sum_{j=3}^{13} \beta_j \ddot{X}_{i,t} + \ddot{u}_{i,t} \qquad (E.3)$$

Where the "trema" above each variable or letter represents its demeaned value. The "$u_{i,t}$" represents the error term, and the remaining notation is carried over from E.1 and E.2. One may also consider these transformed variables to now represent their deviations from their average value in country 'i'. Further note that the general FE specification using tariff rates as the main independent variable of interest was excluded due to issues concerning its validity. This will be further explained in the results section.

Though the interpretations of the estimated coefficients using OLS and FE methods are straightforward and intuitive, they are still vulnerable to a variety of limitations. The most significant being that both approaches are susceptible to issues of reverse causality – similar to many trade-related papers listed in the literature review section. Nonetheless, these static results act as a basic benchmark before moving onto more complex statistical methods.

III. C. i. Dynamic approach

Given the earlier findings in the literature of a characteristic of strong path dependence in ECI (Hidalgo & Hausmann, 2009), it may be useful to consider a dynamic panel data (DPD) model. The advantage of doing so allows us to regress current values of ECI on its own lag, which under static methods, is not suggested. The primary issue in this area of econometrics is the presence of "dynamic panel bias", first presented by Stephen Nickell (1981). In simple terms, a lagged dependent variable on the right-hand side of the equation is said to be correlated with the fixed effects in the error term. This leads to an inflated estimated autoregressive coefficient, and thus a self-constructed endogeneity issue. In order to resolve this issue, the DPD method used in this paper is the difference generalised method of moments (GMM) estimator, or more commonly known as the Arellano-Bond estimator (Arellano & Bond, 1991). The AB model to be estimated, prior to any transformation, can be specified in its most general form as follows:

$$ECI_{i,t} = \sum_{j=1}^{p} \rho_j ECI_{i,t-j} + \beta_1 X_{i,t} + \beta_2 W_{i,t} + v_i + u_{i,t} \qquad (E.4)$$

Where ρ_j are 'p' parameters of lagged ECI to be estimated, the second term represents a vector of strictly exogenous covariates and its respective vector of parameters, the third term represents a vector of predetermined and endogenous covariates along with their respective vector of parameters, v_i are the country-level fixed effects, and $u_{i,t}$ is the error term. Furthermore, v_i and the $u_{i,t}$ are assumed to be independent for each 'i' over all 't', and the $u_{i,t}$ are independently and identically distributed (i.i.d) over the whole sample with variance σ_u^2.

The AB estimator circumvents dynamic panel bias by transforming the data to remove the fixed effects, and using the transformed lagged levels as internal instruments. Here, the lagged levels of the predetermined variables and endogenous variables are used to form GMM-type instruments, whereas the first differences of the strictly exogenous variables are used as standard instruments[11]. However, the question of which variables should be treated as exogenous, predetermined, or endogenous is subjective. In general, exogenous variables are uncorrelated with the past and future values of the error term, i.e. unexplained variances in ECI captured by the error term does not affect its values. If these unexplained variances in ECI affect a variable's value in future periods but not contemporaneously, then it is predetermined. Consequently, if a variable is affected by unexplained variances in ECI contemporaneously, it is endogenous. The following paragraph will outline the rationale for the categorisation of the different variables used in this study.

Conceptually, it would be problematic to categorise almost any of the variables used here as strictly exogenous. For example, other important macroeconomic variables such as public debt levels, exchange rate volatility, or public education expenditure may have an impact on both a country's ECI as well as on its covariates. Thus, it is may be more appropriate to focus on whether covariates are contemporaneously correlated to the error term, or not. This study chooses to treat total population, service sector size, and the six governance indicators as predetermined variables. For example, higher expenditures on public education may positively affect the capabilities available in an economy and thus positively impact its ECI. However, any effect that education could have on the population, the service sector, and even the governance indicators are unlikely to take place immediately. On the other hand, GDP per capita, exports, imports, applied tariffs, and FDI inflows and outflows, are more likely to experience a more contemporaneous relationship for ECI. For example, aggregate country-level productivity improvements in a certain year may lead to a higher level of both ECI and GDP per capita. Similarly, treating trade openness and policy as endogenous help account for reverse causal relationships.

[11] The difference between the two types of instruments lies in their contribution to the instrument matrix. More specifically, GMM-type instruments uses the lags of a variable to contribute multiple columns for each lag to the resulting instrument matrix, whereas each standard instrument contributes only one column to the matrix.

Having set up the general framework for the dynamic methodology section of this paper, it is imperative to discuss some important econometric details of the Arellano-Bond estimator. One of the most significant issues is that the matrix of instruments produced by the AB approach can become unmanageably large with more time periods. As outlined by Roodman (2009), a large instrument collection can "overfit" endogenous variables, leading to biased estimates. Furthermore, time period dummies were generated and included in the AB specification. Since the autocorrelation test and robust estimates of the coefficients' standard errors assume no correlation across individuals in the error term, the inclusion of time dummies helps absorb period specific effects that may homogeneously affect all countries. Time dummies are included in the AB specification as the sole exogenous variables as follows:

$$ECI_{i,t} = \sum_{j=1}^{p} \rho_j ECI_{i,t-j} + \beta_1 W_{i,t} + \beta_2 Year\ dummy_{i,t} + v_i + u_{i,t} \qquad \text{(E.5)}$$

Where the third term now represents a vector of period specific dummy variables – treated as exogenous variables, and the remainder of the specification is identical to that of E.4. Equation 5 is thus the general AB specification to be estimated using the proprietor Stata programme "xtabond2" (Roodman, 2006). Furthermore, Arellano-Bond autoregressive (AR) tests for autocorrelation of orders 1 and 2, the Sargan test, and the Hansen test of over-identified restrictions are all used as diagnostic tests for the purpose of testing instrument validity. Note that the Sargan test is not robust to non-spherical errors – i.e. inconsistent, but is not weakened by many instruments. The Hansen test, on the other hand, is robust to non-sphericity in the errors but is weakened by many instruments. Note that because of the Hansen test's characteristics it is used as the primary test for favourably discriminating between models.

This final paragraph takes note of important decisions made when estimating difference GMM in the following section. Firstly, as mentioned earlier, instrument proliferation can create a significant problem in overfitting the model. There are three main approaches that this research uses to get around this issue. The first is to reduce the time dimension by collapsing the dataset to give two-year averages for each variable. The resulting time span remains the same, from 1996 until 2015, thereby retaining information across the entire sample but measures time in terms of two-year periods. Secondly, due to a large number of covariates used in this study, it is necessary in some cases to collapse the GMM-style instrument matrix – as suggested by Roodman (2006). However, this leads to a reduction in the amount of information available for use in estimation. The third approach is to limit the number of lags that the GMM-style instruments can use, which similarly reduces the amount of information available. These three approaches will be used in various combinations in order to ensure that the number of instruments generated is always less than or equal to the number of panels included in estimation. Next, this study chooses to use two-step GMM estimation rather than its one-step alternative. Due to the introduction of the Windmeijer (2005) small sample correction for standard errors, two-step GMM is said to perform slightly better in estimating coefficients with lower bias and standard errors. Thus, all the following two-step GMM standard errors make use of this Windmeijer correction. Finally, since the panel dataset of this study is unbalanced in terms of observations across all covariates, the forthcoming estimations are based on the forward orthogonal deviations transformations rather than first-difference transformations. The main purpose of doing so is to limit data loss in the cases of gaps in yearly observations for some variables.

III. D. Preliminary Analyses

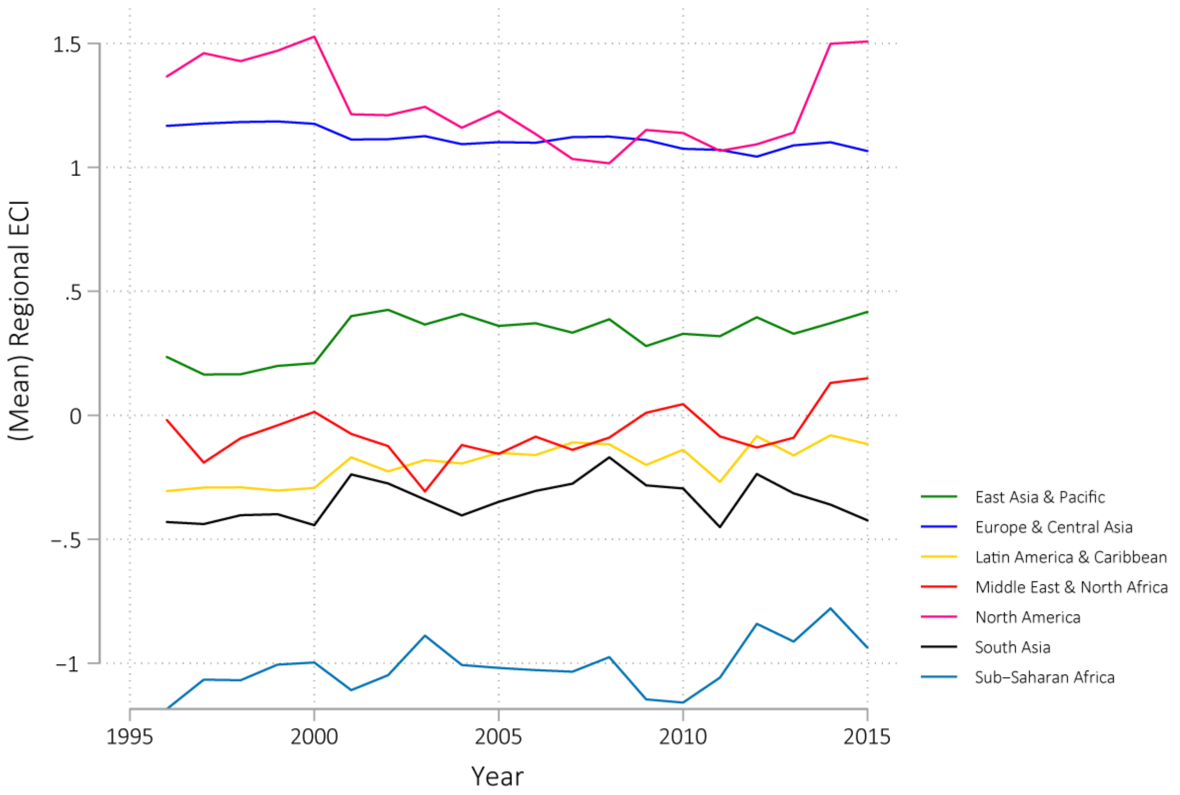
In order to obtain some preliminary insight into the relationships of the covariates with ECI, correlation coefficients are used to measure their direction and strength. A list of all the covariates' correlation with ECI is listed in table A4 in the appendix. First focusing on the relationship between ECI and the trade measures, it is clear that the measures of trade value share a stronger relationship with economic complexity than tariffs. Indeed, it seems that our chosen measure of tariffs shares a weak relationship with the complexity of an economy at first sight – though its sign meets our expectations. Table A5 shows the correlation coefficients between the measures of trade themselves. Again, it appears that despite their importance as an instrument of trade policy, tariffs share a weak negative relationship with total trade values. Returning to Table A4, ECI correlates quite strongly with the log-level value of GDP per capita in accordance with the findings of Hidalgo and Hausmann (2009). Total population, on the other hand, shares the weakest correlation in comparison to the other variables. Service sector size interestingly shares a strong correlation with ECI, justifying our earlier remarks on its relevance in this study. FDI inflows and outflows together share a rather moderate relationship with ECI, and finally, the governance indicators correlate quite variably. The strongest governance correlation with ECI is with government effectiveness, whereas the weakest correlation is with political stability. Although it is not suggested to use these results as the basis of any real economic conclusions, these results do show that all covariates (except for tariffs) are, in fact, positively correlated with ECI.

Another useful form of preliminary analysis is to view how ECI changes over the sample period. Figure 1 displays the development of ECI over time for different regions in the world. The categorisation of each country into a specific group is based on their regional classification according to the World Bank[12] and is also detailed in table A3. It is certainly possible to pursue other forms of categorisation based on interesting factors such as World Bank lending category, income group, cultural similarity, common language, etc. However, these options are better left to later discussions and are thus out of this paper's research scope.

Figure 1 overlays the average regional ECI over the period 1996 until 2015. Other than the clear differences in average ECI for each region, the most relevant insight is that average ECI does not appear to follow any strong trend over time for any region. For example, the average ECI for Europe and Central Asia is the most stable, with an average ECI only marginally lower in 2015 than its value in 1996. Although there is some evidence of weak positive trends for the East Asia and Pacific, Middle East and North Africa, Latin America and Caribbean, and Sub-Saharan Africa regions, the end of period values of average ECI is not drastically different from their starting values. There is also a higher degree of volatility in average ECI for most regions, except for Europe and Central Asia. One explanation for this difference in volatility is the number of countries allocated to each category. Europe and Central Asia, for instance, has the highest number of countries allocated to it (20), whereas one of the most volatile region – North America – only consists of two countries, the United States and Canada. Another contributing explanation is that the European and Central Asian region contains countries that are more stable in ECI. For example, the Latin America and Caribbean region also has a high number of countries in its category (18) but is still substantially more volatile in comparison to Europe.

[12] https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

Figure 1: Mean regional Economic Complexity Index *[ECI]* by year



It may also be relevant to look at how both types of trade measures develop over time. In the case of our measures of trade value, both imports and exports are summed up to create the measure of total trade value or flows. This makes helps avoid too many overlapping lines and thus a more interpretable graph. Figures 2 and 3 overlay the average natural log of imports plus exports and the average weighted-mean applied tariffs, respectively, for each region by year. Figure 2 clearly shows an upward trend for the natural log of imports plus exports over time across all regions. This is to be expected given recent trends in globalisation and trade. The relative positions of each region also follow closely with common knowledge of trade patterns across the world, where the North American trend is clearly driven by the trade flows of the United States. Another notable detail is the visual lack of volatility in values, which is likely due to the natural log transformation of the trade flows. Nonetheless, there are still subtle indications of volatility. For example, the years of 2008 and 2009 suffer from a slight decrease in trade flows – a clear reference to the impact of the 2008 financial crisis on trade.

Figure 2: Mean regional log trade flow values by year

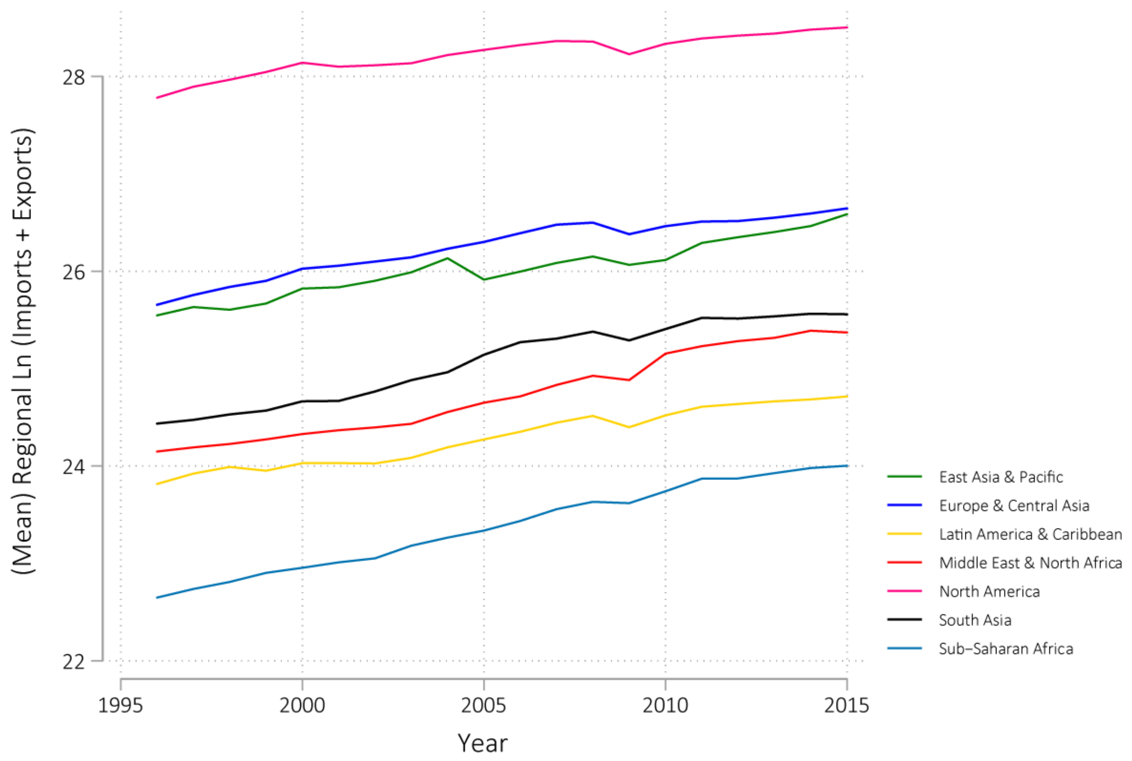


Figure 3: Mean regional weighted-mean applied Tariffs *[TAWM]* by year

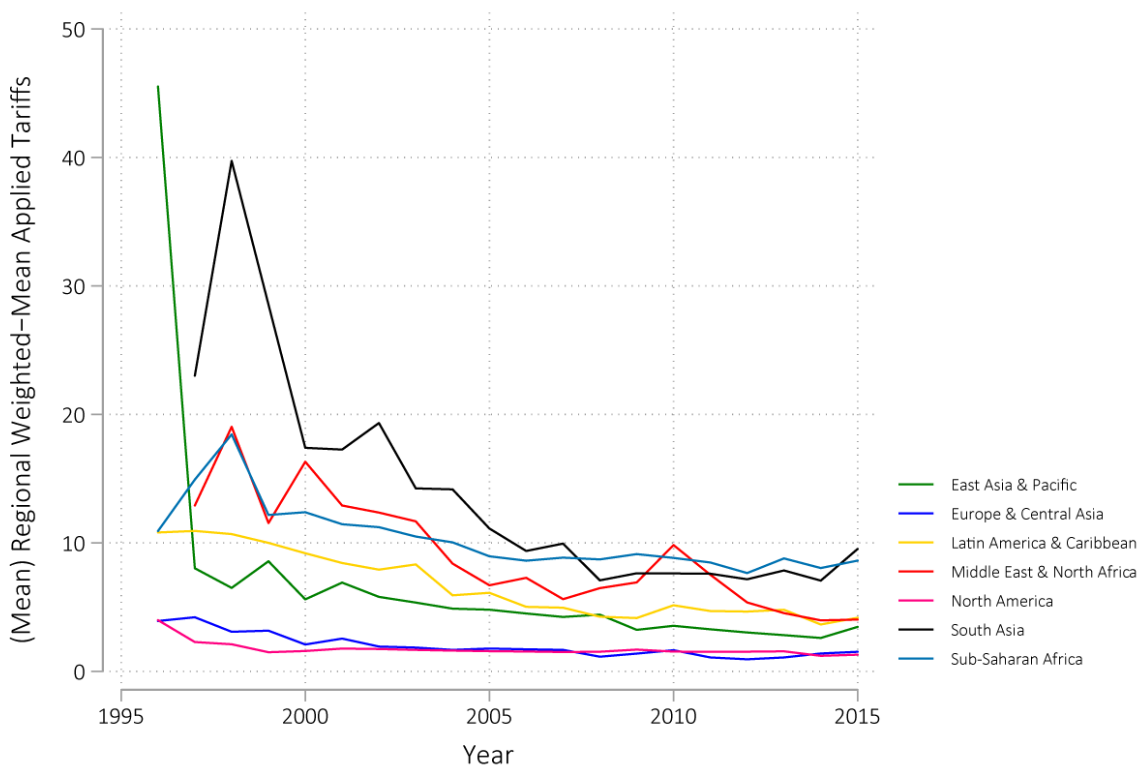


Figure 3 on the other hand, is quite the opposite. Average weighted-mean applied tariff rates are quite volatile over time, save for those of the Europe & Central Asian and North American regions. This difference is likely due to the fact that many of the countries that belong to these regions have already been a part of multilateral agreements such as the General Agreement on Tariffs and Trade (GATT), and have been important founding members of the WTO. In general, figure 3 shows a downward trend across all regions over time. The most significant reductions in weighted-mean tariffs appear to have taken place in the East Asia & Pacific and South Asian regions. It appears that the source of the rather astronomical average tariff value at the start of the period, for the East Asian and Pacific region, is attributable to Japan's weighted-mean applied tariff of 254.58% in 1996. It is unclear why such a high value came to be when the values of the same measure in 1995 and 1997 were 4.97% and 2.95% respectively. It might be concluded that this outlier is the result of either (1) a highly unordinary tariff barrier in the year 1996, or (2) a data entry error. The volatility in the late 1990's for the South Asian region, however, appears to be a result of missing data across countries. More specifically, the data for Pakistan, India, and Sri Lanka (the only three countries in the region), only has data for Pakistan in the year 1998 (39.73%) where India and Sri Lanka's data are missing. On the other hand, the data in 1997 only has data for India and Sri Lanka (20.13% and 25.90% respectively). Quite clearly, the average increase in tariff rates from 1997 to 1998 appears to be much stronger than it may actually be. Thus, results based on this tariff data must be closely examined to ensure that they are not driven by anomalies.

IV – Results

V.A. Static results

We begin with the results obtained from OLS estimation. Table A6, in the appendix, includes four different model specifications, all with the measures of trade value as the main independent variables and each differing in the type of covariates included included. Each model adds a set of control variables in the order of macroeconomic controls, FDI flows, and finally governance indicators. Again, it is important to stress that because the number of observations falls as we add more covariates, one must be cautious when interpreting the results.

The most prominent pattern in table A6 is the robust significance of the log value of imports at the 1% level across all specifications. In contrast, the log value of exports seems consistently unimportant and even negative across all specifications. The coefficients of the control variables also provide some interesting insight, where both the natural logs of GDP per capita and total population are insignificant across all specifications. Service sector size, on the other hand, is significant at the 5% level under OLS models II and III and at the 10% level under OLS model IV. FDI inflows are significant at the 5% level under models III and IV with a small negative coefficient. FDI outflows, however, is insignificant. Finally, the governance controls are generally insignificant save for the indicator for government effectiveness.

Table A7 presents the results in a similar fashion for the weighted-mean applied tariffs. Since the value of weighted-mean applied tariffs in 1996 for Japan is un-ordinarily high – as identified in figure 3 – it is excluded from the set of observations prior to estimation. Table A7 clearly shows a robust, significant, and negative coefficient at the 1% level for OLS model V, and at the 5% level for the remaining models (VI, VII, and VIII). The result is robust to both the inclusion of additional controls, as well as a reduction in observations. However, upon the addition of the macroeconomic controls, the magnitude of the coefficient falls from a relatively high level of -0.1099 to -0.0182. All macroeconomic controls are significant at the 1% level across all the specifications. Notably, both measures of FDI flows are insignificant across models VII and VIII and the governance indicators' coefficient has similarly experienced some stark changes. Under model VIII, three governance indicators have significant coefficients. Government effectiveness is significant at the 1% level, whereas the coefficients for control of corruption and political stability are significant at the 5% level.

Turning to the FE estimation results, table A8 presents the results of the FE models with the trade value measures as the main independent variables of interest. Firstly, it appears that the inclusion of fixed effects greatly reduces the importance of trade value with respect to economic complexity. Although it retains a positive coefficient significant at the 5% level under FE model I, its significance disappears as we add the other covariates. Furthermore, the magnitude of the coefficient is greatly reduced from its value of 0.6527 under OLS model I to 0.1365 under FE model I. Since the number of observations is identical between these two specifications, it is clear that the fixed effects have absorbed much of the explanatory power previously attributed to the log value of imports in OLS model I. FE model II's coefficients of the macroeconomic controls with respect to its OLS counterpart (model II), have also changed significantly. Although natural log of GDP per capita coefficient is insignificant across all the FE specifications, the coefficient of the natural log of total population gains importance and magnitude. Although it is significant at the 5% level under FE model II, its significance at the 10% level under FE models III and IV implies that the addition of the other covariates diminishes its explanatory power. Service sector size instead continues to remain robust at the 5% level across all of the FE specifications, albeit with a lower magnitude than under the OLS specifications. FDI flows have also become insignificant under FE models III and IV, corroborating our findings under the tariff rates. It is therefore evident that FDI flows in general share a weak relationship with respect to ECI when cross-country differences are accounted for. Finally, the governance indicators retain the same pattern as the results under table A6, where only government effectiveness hold a significant coefficient at the 5% level. Subsequently running Hausman tests (Table A9) for the FE results to test them against the results of their corresponding random effects (RE) specifications shows that the FE model is unequivocally preferred. As mentioned earlier, the FE results with weighted-mean applied tariffs appear to be invalid across all specifications and are therefore excluded from this analysis. More specifically, results of all its specifications' F-tests return p-values that are insignificant.

Prior to moving onto the dynamic results, it would be useful to observe whether or not there exists lagged effects that may have been overlooked in prior estimations. Table A10 shows estimation results for the full specifications of both sets of OLS models, as well as the FE model for import and export values. Note that the FE model for tariff rates remains invalid, with insignificant p values for its F-test.

The results for OLS model IX is best comparable to OLS model IV. The most interesting discovery is that the coefficient estimates for both the current values of log imports and exports are now insignificant, whereas the coefficients of the first lags of both log imports and exports are both significant – at the 5% and 1% level respectively. Furthermore, service sector size becomes insignificant under model IX, whereas both FDI inflows and government effectiveness retain the magnitude, sign, and significance of their coefficients. As for the tariff results, OLS model X is best compared with OLS model VIII. Here, the results are largely unchanged with no evidence of a lagged effect. The remaining coefficients for all variables do not change much in terms of magnitude, with values close to their counterparts under OLS model VIII. Finally, the FE model V is best compared to FE model IV. Interestingly, the results for the current levels of both imports and exports are both insignificant, echoing prior outcomes. However, the first lag of log exports now becomes significant at the 1% level, whereas the first lag of lag imports is insignificant. The outcomes for the other coefficients are mostly similar, however, the estimate for total population is no longer significant. Similarly, the coefficient of government effectiveness is now significant at the 10% level under FE model V. The only coefficient that is (almost) identical between the two models is that of service sector size. In conclusion, it appears that accounting for lagged effects is more relevant for specifications using imports and exports – allowing for the identification of new insights.

Table A6 – Ordinary Least Squares (OLS) Results – Trade Value [*ln (Imports) & ln (Exports)*]

Dependent Variable Economic Complexity	**OLS Model I**		**OLS Model II**		**OLS Model III**		**OLS Model IV**	
Independent Variables	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**
Constant	-11.2721***	(0.7170)	-9.0681***	(0.5814)	-9.8169***	(0.6088)	-7.8476***	(0.9334)
Ln (Imports)	0.6527***	(0.1491)	0.6061***	(0.1029)	0.5880***	(0.1243)	0.5047***	(0.1297)
Ln (Exports)	-0.1847	(0.1423)	-0.2706*	(0.1370)	-0.1905	(0.1659)	-0.1760	(0.1590)
Ln (GDP per capita)			0.1947	(0.1390)	0.1566	(0.1493)	0.0506	(0.1359)
Ln (Total Population)			-0.0934	(0.0971)	-0.1203	(0.1012)	-0.0687	(0.1063)
Service Sector Size			0.0146***	(0.0048)	0.0155***	(0.0053)	0.0094*	(0.0055)
FDI Inflows					-0.0043**	(0.0019)	-0.0039**	(0.0018)
FDI Outflows					0.0021	(0.0015)	0.0020	(0.0015)
Control of Corruption							-0.2114	(0.1281)
Government Effectiveness							0.4608***	(0.1443)
Political Stability							0.1228*	(0.0738)
Rule of Law							-0.1444	(0.1280)
Regulatory Quality							0.0874	(0.1242)
Voice and Accountability							0.0931	(0.0725)
Observations	1,540		1,444		1,348		1,169	
Sample Period	1996 - 2015		1996 - 2015		1996 - 2015		1996 - 2015	
Probability > F	0.0000		0.0000		0.0000		0.0000	
R^2	0.6207		0.7676		0.7721		0.7988	

*Standard Errors are adjusted for clusters in ISO (at the country-level). Asterisks indicate significance levels of 10%, 5%, and 1% using *, **, and *** respectively. All figures are rounded to four decimal places.*

Table A7 – Ordinary Least Squares (OLS) Results – Tariffs, Applied, Weighted-Mean (*TAWM)*

Dependent Variable Economic Complexity	**OLS Model V**		**OLS Model VI**		**OLS Model VII**		**OLS Model VIII**	
Independent Variables	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**
Constant	0.8588***	(0.1329)	-7.6536***	(0.6815)	-7.9289***	(0.7834)	-5.8761***	(0.8651)
Applied Tariff Rates	-0.1099***	(0.0164)	-0.0182**	(0.0087)	-0.0186**	(0.0091)	-0.0134**	(0.0067)
Ln (GDP per capita)			0.4297***	(0.0457)	0.4436***	(0.0524)	0.2327***	(0.0541)
Ln (Total Population)			0.1766***	(0.0324)	0.1884***	(0.0356)	0.1877***	(0.0364)
Service Sector Size			0.0203***	(0.0045)	0.0200***	(0.0046)	0.0135***	(0.0051)
FDI Inflows					-0.0019	(0.0018)	-0.0024	(0.0014)
FDI Outflows					0.0008	(0.0016)	0.0012	(0.0014)
Control of Corruption							-0.3939**	(0.1590)
Government Effectiveness							0.6650***	(0.1604)
Political Stability							0.1820**	(0.0730)
Rule of Law							-0.0652	(0.1452)
Regulatory Quality							0.1675	(0.1348)
Voice and Accountability							-0.0486	(0.0784)
Observations	1,390		1,307		1,238		1,087	
Sample Period	1996 - 2015		1996 - 2015		1996 - 2015		1996 - 2015	
Probability > F	0.0000		0.0000		0.0000		0.0000	
R^2	0.3066		0.7292		0.7211		0.7753	

*Standard Errors are adjusted for clusters in ISO (at the country-level). Asterisks indicate significance levels of 10%, 5%, and 1% using *, **, and *** respectively. All figures are rounded to four decimal places.*

Table A8 – Within Fixed Effects (FE) Results – Trade Value [*ln (Imports) & ln (Exports)*]

Dependent Variable Economic Complexity	**FE Model I**		**FE Model II**		**FE Model III**		**FE Model IV**	
Independent Variables	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**
Constant	-1.3827	(0.9418)	-5.3114**	(2.1816)	-5.7597**	(2.2982)	-5.1504***	(1.9442)
Ln (Imports)	0.1365**	(0.0610)	0.0606	(0.0673)	-0.0162	(0.0772)	0.0451	(0.0762)
Ln (Exports)	-0.0732	(0.0745)	-0.1394	(0.0982)	-0.0254	(0.0964)	-0.0557	(0.0992)
Ln (GDP per capita)			0.1047	(0.1513)	0.0922	(0.1936)	0.0228	(0.1924)
Ln (Total Population)			0.3577**	(0.1536)	0.3380*	(0.1819)	0.2924*	(0.1544)
Service Sector Size			0.0087***	(0.0032)	0.0096**	(0.0041)	0.0087**	(0.0038)
FDI Inflows					-0.0003	(0.0004)	-0.0001	(0.0003)
FDI Outflows					-0.0002	(0.0005)	-0.0002	(0.0004)
Control of Corruption							0.0297	(0.0544)
Government Effectiveness							0.1757**	(0.0750)
Political Stability							-0.0316	(0.0354)
Rule of Law							-0.0849	(0.0853)
Regulatory Quality							0.0134	(0.0558)
Voice and Accountability							-0.0305	(0.0656)
Observations	1,540		1,444		1,348		1,169	
Sample Period	1996 - 2015		1996 - 2015		1996 - 2015		1996 - 2015	
Probability > F	0.0195		0.0350		0.0126		0.0053	
R^2 within	0.0203		0.0489		0.0571		0.0652	
R^2 between	0.6740		0.1171		0.1758		0.2195	
R^2 overall	0.6028		0.0919		0.1456		0.1932	

*Standard Errors are adjusted for clusters in ISO (at the country-level). Asterisks indicate significance levels of 10%, 5%, and 1% using *, **, and *** respectively. All figures are rounded to four decimal places.*

Table A10: Static models with lagged independent variables

Dependent Variable Economic Complexity	**OLS Model V** Imports and Exports		**OLS Model VI** Tariff Rates		**FE Model V** Imports and Exports	
Independent Variables	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**
Constant	-7.9344***	(0.9459)	-5.8346***	(0.8829)	-4.0601*	(2.1100)
Ln (Imports)	0.1442	(0.2565)			0.0269	(0.1007)
Ln (Imports) (t – 1)	0.4354**	(0.2089)			0.0803	(0.0878)
Ln (Exports)	0.4669	(0.3116)			0.1195	(0.0896)
Ln (Exports) (t – 1)	-0.6987***	(0.2534)			-0.2757***	(0.0978)
Applied Tariffs			-0.0145*	(0.0079)		
Applied Tariffs (t – 1)			-0.0023	(0.0077)		
Ln (GDP per capita)	0.0501	(0.1405)	0.2332***	(0.0554)	0.1228	(0.2104)
Ln (Total Population)	-0.0885	(0.1116)	0.1896***	(0.0371)	0.2319	(0.1658)
Service Sector Size	0.0082	(0.0057)	0.0125**	(0.0053)	0.0083**	(0.0040)
FDI Inflows	-0.0043**	(0.0019)	-0.0025*	(0.0015)	-0.0001	(0.0003)
FDI Outflows	0.0022	(0.0016)	0.0014	(0.0014)	-0.0002	(0.0004)
Control of Corruption	-0.1983	(0.1318)	-0.4052**	(0.1637)	0.0248	(0.0566)
Government Effectiveness	0.4697***	(0.1438)	0.6630***	(0.1618)	0.1327*	(0.0737)
Political Stability	0.1087	(0.0735)	0.1803**	(0.0712)	-0.0292	(0.0361)
Rule of Law	-0.1745	(0.1302)	-0.0843	(0.1501)	-0.0987	(0.0799)
Regulatory Quality	0.0855	(0.1262)	0.1970	(0.1440)	0.0007	(0.0565)
Voice and Accountability	0.1019	(0.0754)	-0.0561	(0.0799)	-0.0193	(0.0660)
Observations	1,106		988		1,106	
Sample Period	1996 – 2015		1996 – 2015		1996 – 2015	
Probability > F	0.0000		0.0000		0.0021	
R^2	0.8035		0.7750		0.3208	

*Standard Errors are adjusted for clusters in ISO (at the country-level). Asterisks indicate significance levels of 10%, 5%, and 1% using *, **, and *** respectively. All figures are rounded to four decimal places. Note that the R^2 of FE model V refers to the overall R^2 of the model.*

V.B. Dynamic results

The AB estimation results for the trade flow values are presented in Table A11. AB models I and II differ in their approach towards limiting instrument proliferation. AB model I uses data on two-year averages for each variable, thereby allowing it to use up to 4 lags for GMM-style instruments. AB model II, in contrast, uses the original data set with the current levels for each variable. However, this approach requires a more stringent lag limit of 2. Before comparing outcomes, it would be useful to first take note of each model's p-values with respect to their Sargan and Hansen statistics. Notably, both models' instrument sets appear to be valid under the Hansen test, yet the Sargan test suggests that AB model I may not have a valid instrument set.

Under both models, the autoregressive coefficient of ECI is significant at the 1% level, where the magnitude is higher under AB model II. Importantly, the resulting coefficients of the natural log of imports and exports are remarkably different under both models. AB model I finds insignificant coefficients across both imports and exports for current levels and lagged levels. AB model II, however, finds significant results at the 5% level for the current and first lag level of imports – where the signs are negative and positive, respectively – and significant coefficients at the 10% level for the current and first lag level of exports – where the signs are positive and negative, respectively. Other notable findings are that service sector size's coefficient is significant at the 1% level under AB model I, but is insignificant under AB model II. Furthermore, the coefficients of rule of law are significant at the 5% level under AB model I, but not significant under AB model II. The remaining coefficients for the other covariates are insignificant under both models.

The results for AB estimation with tariff rates are available under table A12, where the structure of analysis is identical to that of table A11. Note that the Sargan test for both models indicates instrument invalidity, where both statistics are significant at the 5% level. Furthermore, the Hansen J-statistics' p-values under AB model IV is also significant at the 5% level, whereas under AB model III is insignificant. Therefore, the only potentially valid outcomes are under AB model III since the Hansen test is used as the primary discriminatory diagnostic test. AB model III's autoregressive coefficient for ECI is significant at the 1% level, with a magnitude close to AB model I's estimate. However, as for the tariff rates, both the first lag and current value coefficients are insignificant in contradiction to prior results under OLS estimation. Other noteworthy outcomes are that FDI inflows, FDI outflows, total population, and service sector size are significant at the 10% level.

Table A11 – Arellano-Bond two-step (difference) GMM Results:
Trade Value [*ln (Imports) & ln (Exports)*]

Dependent Variable Economic Complexity	**AB Model I**		**AB Model II**	
Independent Variables	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**
Economic Complexity (t – 1)	0.4411***	(0.1188)	0.5999***	(0.1480)
Endogenous variables				
Ln (Imports)	0.3683	(0.3298)	-1.3344**	(0.5190)
Ln (Imports) (t – 1)	0.0487	(0.2402)	1.1853**	(0.5000)
Ln (Exports)	-0.2000	(0.4600)	1.6565*	(0.8483)
Ln (Exports) (t – 1)	-0.1587	(0.3299)	-1.4433*	(0.8161)
Ln (GDP per capita)	-0.3421	(0.4995)	-0.1521	(0.6032)
FDI Inflows	0.0005	(0.0018)	0.0003	(0.0023)
FDI Outflows	-0.0005	(0.0016)	-0.0016	(0.0030)
Predetermined variables				
Ln (Total Population)	0.3348	(0.3014)	0.2168	(0.3242)
Service Sector Size	0.0296***	(0.0103)	0.0108	(0.0126)
Control of Corruption	0.1028	(0.1763)	0.2026	(0.2878)
Government Effectiveness	-0.1059	(0.1804)	-0.1972	(0.1984)
Political Stability	0.0552	(0.1096)	-0.0404	(0.1117)
Rule of Law	-0.2810**	(0.1407)	0.0628	(0.2178)
Regulatory Quality	0.0343	(0.1569)	0.1651	(0.1429)
Voice and Accountability	-0.2555	(0.1800)	0.0010	(0.1779)
Period/Year dummies	[Results Omitted]		[Results Omitted]	
Observations	546		1028	
Sample Period	1996 – 2015 [2 year averages]		1996 – 2015 [Current Levels]	
Number of groups	78		78	
Instruments	64		43	
GMM-style instruments lag limit	4		2	
Sargan test statistic [prob. > chi2]	0.002		0.373	
Hansen J-statistic [prob. > chi2]	0.267		0.468	
AR (1) test [prob. > z]	0.008		0.000	
AR (2) test [prob. > z]	0.757		0.425	

*Two-step difference GMM standard errors are robust to heteroskedasticity and autocorrelation within individuals, and are adjusted according to the Windmeijer (2005) correction. Asterisks indicate significance levels of 10%, 5%, and 1% using *, **, and *** respectively. All figures are rounded to four decimal places. Dummy variables for each year were estimated in both models, however, their respective coefficient estimates are omitted due to their irrelevance to the research question and to conserve space. Lag limits for the GMM-style instruments are set to a minimum of one lag and the maximum lag is specified. Data transformation under difference GMM is conducted under the forward orthogonal deviations transform.*

Table A12 – Arellano-Bond two-step (difference) GMM Results:
Tariffs, Applied, Weighted-Mean (*TAWM)*

Dependent Variable Economic Complexity	**AB Model III**		**AB Model IV**	
Independent Variables	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**
Economic Complexity (t – 1)	0.3633*	(0.1832)	0.3723***	(0.1115)
<u>Endogenous variables</u>				
Applied Tariff Rates	0.0109	(0.0189)	-0.0466	(0.0363)
Applied Tariff Rates (t – 1)	-0.0092	(0.0104)	0.0408	(0.0418)
Ln (GDP per capita)	0.2594	(0.3365)	-0.0643	(0.4035)
FDI Inflows	0.0030	(0.0017)	0.0008	(0.0015)
FDI Outflows	-0.0025	(0.0014)	-0.0011	(0.0019)
<u>Predetermined variables</u>				
Ln (Total Population)	0.5349	(0.2524)	0.3729	(0.2979)
Service Sector Size	0.0283	(0.0156)	0.0205	(0.0208)
Control of Corruption	0.0627	(0.1364)	0.3631	(0.3528)
Government Effectiveness	0.0960	(0.1645)	-0.0391	(0.2357)
Political Stability	-0.0473	(0.1167)	-0.0311	(0.1889)
Rule of Law	-0.2127	(0.2305)	0.1648	(0.3925)
Regulatory Quality	0.0841	(0.1618)	0.0776	(0.2112)
Voice and Accountability	-0.1312	(0.1535)	0.0317	(0.4229)
Period/Year dummies	[Results Omitted]		[Results Omitted]	
Observations	520		908	
Sample Period	1996 – 2015 [2 year averages]		1996 – 2015 [Current Levels]	
Number of groups	79		80	
Instruments	60		41	
GMM-style instruments lag limit	4		2	
Sargan test statistics [prob. > chi2]	0.045		0.001	
Hansen J-statistic [prob. > chi2]	0.109		0.031	
AR (1) test [prob. > z]	0.017		0.000	
AR (2) test [prob. > z]	0.352		0.197	

*Two-step difference GMM standard errors are robust to heteroskedasticity and autocorrelation within individuals, and are adjusted according to the Windmeijer (2005) correction. Asterisks indicate significance levels of 10%, 5%, and 1% using *, **, and *** respectively. All figures are rounded to four decimal places. Dummy variables for each year were estimated in both models, however, their respective coefficient estimates are omitted due to their irrelevance to the research question and to conserve space. Lag limits for the GMM-style instruments are set to a minimum of one lag and the maximum lag is specified. Data transformation under difference GMM is conducted under the forward orthogonal deviations transform.*

It may be noted that it is possible to limit instrument proliferation by another approach. That is, by limiting the number of covariates included in the specification. In this case, only the main independent variables of interest are used to give parsimonious specifications. This approach allows estimation to take place without the need to limit the number of lags used by

GMM instrumentation or the need to collapse the instrument matrix. However, these results are qualitatively identical to findings under AB models I and III and provide no additional insights to what was already found. Therefore in the interest of conserving space, they are omitted from the following discussions.

The following section will discuss the results as a whole, reflecting on the most relevant findings as well as discussing potential reasons for particular outcomes. Furthermore, discussions on the limitations of this research will be outlined as a reference for future research.

V – Discussion and Limitations

Focusing first on the results for the log values of imports and exports, the overall finding is that the relationship between direct trade openness and ECI is quite convoluted. With respect to the first set of static econometric approaches (OLS models I – IV & FE model I – IV), it appears that the coefficient estimates of the current level of log imports are likely to be driven by its own relationship with some unobserved, time-invariant cross-country differences. The current level of log exports, however, are found to be unimportant from the get-go – indicating that its explanatory power is largely overshadowed by its counterpart, log imports. Turning to the outcomes for the lagged levels (OLS model IX and FE model V), the proposition above on log imports' relationship with unobserved fixed effect(s) is corroborated. Specifically, the lagged value of log imports becomes insignificant upon the inclusion of country specific fixed effects. In contrast, the lagged value of log exports is significant at the 1% level under both models. This finding shows that prior specifications somehow obscured the existence of a lagged relationship between exports and ECI. Under OLS model IX, a 10% increase in imports for a given year is, on average, associated with an approximate 0.04354 point increase in ECI the following year, ceteris paribus. Likewise, a 10% increase in exports for that same year is, on average, associated with an approximate 0.06987 point decrease in ECI the following year, ceteris paribus. Under FE model V, only the export effect remains significant – albeit with a smaller associated effect. Thus, a 10% increase in exports for a given year, on average, is associated with an approximate 0.02757 point decrease in ECI, ceteris paribus. These lagged export effects appear counterintuitive, however, could be explained by adverse effects of specialisation on the productive structure of a country. For example, effects associated with Dutch disease. The positive lagged effect associated with imports, seems to point towards the benefits of importing goods. For example, technology and knowledge transfers associated with importing goods. However, this effect is deemed unimportant when accounting for time invariant cross country differences.

The dynamic approach results are more difficult to interpret. The first complication is the stark difference between AB model I and II. As a reminder, AB model I is based on data that collapses variables into two-year averages in an attempt to control instrument proliferation. Although AB model I makes use of more information – substantiated by its larger instrument count – it may be based on data that is not particularly reflective of reality. Furthermore, AB model I's Sargan statistic results in a p-value that suggests that its instruments are jointly insignificant. This is important as although the Sargan test is the Hansen test's non-robust counterpart, it is not weakened by many instruments. Therefore, it may be more appropriate to refer to AB model II's results. Here, the coefficients of imports – both lagged and current levels – are significant at the 5% level, whereas that of exports are significant at the 10% level. This mirrors earlier findings under OLS model IX and FE model V, although in the case of AB model II, current levels are also found to be important. The most intriguing finding are the signs of the coefficients – suggesting a rather complex relationship between trade and ECI. For instance, holding all other factors constant, a one-time increase in current imports of 10% is associated with a contemporaneous decrease of approximately 0.1334 points in ECI. The next year, however, the effects are carried forward by the autoregressive coefficient of ECI as well as in the lagged effect of the same increase in imports. Therefore, the following year experiences a net increase in ECI of approximately 0.0385[13] points, leading a net decrease over the two years of 0.0950 points in ECI, ceteris paribus. In contrast, the associated effects with respect to exports are reversed. Using the same exercise, a one-time 10% increase in current exports, ceteris paribus, is associated with a contemporaneous increase of approximately 0.1657 points in ECI. The following year subsequently experiences a net decrease of approximately 0.0450[14] points in ECI, ceteris paribus, as a result of the lagged export effect combined with the autoregressive coefficient of ECI from the previous period. The net increase in ECI over the two years is therefore approximately 0.1207 points.

Naturally, it would be a fallacy to consider the magnitudes of these effects to be exact representations of the true relationship between trade and economic complexity. However, AB model II does broadly point towards the existence of countervailing effects of trade in relation to economic complexity. For example, the contemporaneous negative effect of imports may be

[13] 0.11853 – (0.5999 * (-0.13344)) = 0.038479
[14] (0.5999 * 0.16565) – 0.14433 = - 0.04496

associated with immediate adverse import-competing effects on local firms, whereas its lagged positive effect may be associated with productivity, knowledge, or technological benefits that materialise later. On the other hand, the contemporaneous positive effect of exports is likely to be a result of the direct benefits of exporting – i.e. industry growth, whereas, as mentioned in the previous paragraph, the negative lagged effect may be related to Dutch disease-type effects due to specialisation. Interestingly, Dutch disease was noted by Hidalgo (2009, p. 12) in a discussion on the possible reasons for relative reductions in ECI.

Now to summarise the different results for weighted average applied tariffs. The results under the OLS specifications tend to indicate a negative association between tariff rates and ECI. Indeed, this association remains robust to the inclusion of various controls, as well as a reduction in the number of observations. Under OLS model VIII, the full specification for tariff rates, a 10%-point increase in weighted average applied tariff rates leads to, on average, an approximate 0.134 point decrease in ECI, ceteris paribus. Attempting to identify some additional lagged effect of tariff changes under OLS model X, leads to little change in the magnitude of the coefficient of the current level of applied tariffs. More specifically, the same 10%-point increase as above is associated, on average, with a larger 0.145 point decrease in ECI, ceteris paribus. Furthermore, OLS model X does not identify any lagged effect. The dynamic results for tariff rates under AB models III and IV, however, do not lead to any significant results. It may be concluded that upon accounting for the autoregressive dynamics in ECI, the explanatory power of applied tariff rates are substantially weakened. Although the static approach outcomes find evidence of some relationship, it is likely that the patterns of tariff data contribute to its insignificance under the AB models. For instance, it is useful to compare figures 1 and 3 – earlier presented in the preliminary analysis. If we focus on the ECI and tariff rate patterns for a single region, for instance, North America , tariff rates are substantially less volatile compared to ECI. Average regional tariff rates for North America are, along with those of Europe and Central Asia, lower than all other regions in the sample and are quite stable over time. Interestingly, North America has the highest average levels of ECI compared to other regions – except for Europe and Central Asia – but it is also substantially more volatile than its pattern in tariff rates. For example, between the years 2000 and 2005, average tariff rates in North America did not visibly change. However, in that same period, average ECI had visibly decreased. Clearly, if average tariff rates did not change, how can they provide explanatory power for this change in average ECI? In conclusion, despite the significant results found under the static approach for tariff rates, the insignificant results under a dynamic approach indicate that even if these relationships exist, they are likely to be economically weak.

This discussion brings us to the foremost limitation of this study, proxying for trade liberalisation and openness. The main issue in choosing a viable proxy is that the definition of trade liberalisation and openness is subjective. The choice of weighted mean applied tariff rates and trade value flows in this study was mainly grounded in the need for easily interpretable, convenient, and widely available measures. However, despite the outcomes of this paper's econometric approaches, the findings are quite limited. More specifically, this paper's proxy for trade policy, weighted mean applied tariffs, were found to be unimportant for the development ECI in the context of the sample period. But this does not imply that trade liberalisation is generally unimportant. There are many different factors that determine liberal trade policy – for example, the use of NTB's. Similarly, the finding that trade value flows are relevant for ECI development only scratches at the surface of the relationship between trade openness and ECI. As discussed in the literature review section, trade openness can also be defined in a variety of ways and measures (E.g. Krueger, 1978; Balassa, 1982; Sachs & Warner, 1995), which may lead to different insights compared to those found here. Therefore, it is important to refrain from generalising this paper's findings and to instead use these as a benchmark for further research using alternative measures of trade liberalisation and openness.

An alternative view is that there are also more ways to analyse the relationships using the same measures of trade liberalisation and openness used here. In the interest of keeping within the scope of this paper and avoiding an inordinate discussion, all variables used in the above econometric analyses were kept in levels. Another potential approach would be to transform all variables by first differencing them, thereby allowing for the analysis of changes in trade value flows and tariffs on changes in ECI. Such an approach may lead to alternative findings that are interesting in their own right – whether or not they contribute or oppose the findings of this paper. Taking a step further, it would also be of interest to study the results under long differenced variables. For example, studying the 5-year changes in variables allows for the study of longer-term effects without the need of regressing an excessive number of lagged levels.

Finally, the compatibility of this study's dataset with the Arellano-Bond dynamic estimation approach, as well as the AB approach itself is a notable limitation. For instance, the dynamic results are likely to be contingent on the choice of treating the various covariates as endogenous, predetermined or exogenous. Although the choices of treatment were based on the logical context of this study, it is unclear how the results would be affected if these choices were to be refuted. In addition, the time span of the sample as well as the number of covariates involved in a study in this field force this paper's AB approach to limit instrument proliferation at every turn. Citing previous concerns in the methodology section, some of the measures used here to limit instrument proliferation – collapsing the instrument set and setting lag limits for instrumentation – considerably limit the amount of information available for estimation. The other approach to take two-year averages for two-year periods, thereby collapsing the time dimension in half, makes it difficult to compare outcomes with prior results. In conclusion, it would be of interest to consider alternative DPD models that would not face the same issues in future research. Otherwise, in combination with the suggestion of using substitute trade proxies, it may be an option to use a dataset that better fits the context of the AB approach. For instance, alternative trade proxies typically suffer from limited sample period lengths – which is well suited for the AB methodology.

VI – Conclusive remarks and suggestions for future research

This paper is the metaphorical tip of the iceberg that is the relationship between trade and the economic complexity of nations. With the dataset at hand and the econometric methodologies used in this paper, we were able to analyse, at an aggregate level, a brief outline of this relationship. Although there is evidence for a relationship between economic complexity (ECI) and direct trade openness, measured by the total value of trade, the relationship itself is multifaceted and is likely to be defined by a variety of opposing and reinforcing effects. As discussed in the previous section, it appears that the net relationship between exports and ECI is positive, whereas the net relationship of increasing imports and ECI seems to be a negative association. The underlying net effects that embody the nature of these relationships are further differentiated by a contemporaneous effect and a lagged effect of imports and exports, corroborating the complexity of this relationship. These effects are worthy of further investigation in their own right, and, as outlined in prior section, may embody a number of recognised concepts in trade theory and empirics. For example, import-competition effects on industrial dynamics, firm-level productivity and knowledge gains from trade, industry-level productivity gains from trade, and even Dutch-disease type effects that accrue from trade specialisation. On the other hand, this study also shows that weighted mean applied tariff rates may not play an important role in the development of ECI. In the current state of world affairs with respect to trade, this finding may hold some truth due to the overall decline in tariff rates across the world.

Nonetheless, it is important to maintain caution in interpreting these results as this paper merely serves as a basic benchmark for future research. It is paramount that future studies into the relationship between trade and economic complexity make use alternative measures of both trade liberalisation and openness. For example, measures that account for non-tariff barriers, black market exchange rates, etc. These alternative measures will lend further insight into the nature of the relationship with a level of detail that the measures used here cannot embody. However, other suggestions for future research that don't specifically pertain to the limitations of this paper include further investigation into other determinants of economic complexity. For example, throughout this study, a recurring theme for most estimation outcomes was the importance of government effectiveness and service sector size. Future research could, for example, examine more closely the role of factors related with, but not limited to, institution and geographical factors. Moreover, although it was surmised that FDI is likely to have a significant impact on ECI, its role was found to be consistently unimportant. It may be desirable to consider alternative measures of FDI beyond simple inflows and outflows to determine whether its irrelevance is reflective of FDI's true relationship with ECI, or whether it is of an idiosyncratic nature.

Finally, it is natural ask what these results could mean in the grand scheme of things. In terms of policy advice, I believe that this paper may shed some light into the benefits and costs of trade in the context of productive structure. For instance, the countervailing effects of trade value flows quite simply indicate that policy makers must take into account the net effects of various underlying factors on a country's productive structure when engaging in trade activity. It is important, for example, to consider the adverse effects on industry dynamics when a country imports more, as well as the potential adverse effects on productive structure when it exports too much. Weighing these cons with respect to the benefits that also accrue from trade is vital. Furthermore, the findings for applied tariffs seems to indicate that tariffs as a policy instrument for economic complexity may be increasingly unimportant in the current era – particularly given current patterns in tariff reductions. In conclusion, the answer to this paper's research question of "what is the relationship between trade and economic complexity?" is that it's complicated. Like the plethora of research previously conducted on trade liberalisation or openness on economic growth, future research on economic complexity and trade is likely to be as convoluted and controversial.

VIII – References

Abdon, A., & Felipe, J. (2011). The product space: What does it say about the opportunities for growth and structural transformation of Sub-Saharan Africa?.

Acemoglu, D., Johnson, S., & Robinson, J. A. (2000). The colonial origins of comparative development: An empirical investigation (No. w7771). National bureau of economic research.

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. The review of economic studies, 58(2), 277-297.

Bahar D. (2016, November 4). Diversification or specialization: What is the path to growth and development?. *Brookings Institution*. Retrieved from https://www.brookings.edu/research/

Balassa, B. (1965). Trade liberalisation and "revealed" comparative advantage. The Manchester School, 33(2), 99-123.

Balassa, B. (1982). Development Strategies in Semi-Industrialised Countries. Johns Hopkins Univ. Press, Baltimore.

Boschma, R., & Iammarino, S. (2009). Related variety, trade linkages, and regional growth in Italy. Economic geography, 85(3), 289-311.

Bustos, S., Gomez, C., Hausmann, R., & Hidalgo, C. A. (2012). The dynamics of nestedness predicts the evolution of industrial ecosystems. PloS one, 7(11), e49393.

Caldarelli, G., Cristelli, M., Gabrielli, A., Pietronero, L., Scala, A., & Tacchella, A. (2012). A network analysis of countries' export flows: firm grounds for the building blocks of the economy. PloS one, 7(10), e47278.

Cristelli, M., Gabrielli, A., Tacchella, A., Caldarelli, G., & Pietronero, L. (2013). Measuring the intangibles: A metrics for the economic complexity of countries and products. PloS one, 8(8), e70726.

Cristelli, M., Tacchella, A., & Pietronero, L. (2015). The heterogeneous dynamics of economic complexity. PloS one, 10(2), e0117174.

Dollar, D. (1992). Outward-oriented developing economies really do grow more rapidly: evidence from 95 LDCs, 1976-1985. Economic development and cultural change, 40(3), 523-544.

Dollar, D., & Kraay, A. (2003). Institutions, trade, and growth. Journal of monetary economics, 50(1), 133-162.

Edwards, S. (1998). Openness, productivity and growth: what do we really know?. The economic journal, 108(447), 383-398.

Felipe, J. (2012). Inclusive Growth, Full Employment, and Structural Change: Implications and Policies for Developing Asia. Anthem Press.

Frankel, J. A., & Romer, D. (1999). Does trade cause growth?. American economic review, 379-399.

Freitas, E. E., & Paiva, E. A. (2016). Diversificação e sofisticação das exportações: uma aplicação do product space aos dados do Brasil. Revista Econômica do Nordeste, 46(3), 79-98.

Frenken, K., Van Oort, F., & Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. Regional studies, 41(5), 685-697.

Greenaway, D., Morgan, W., & Wright, P. (2002). Trade liberalisation and growth in developing countries. Journal of development economics, 67(1), 229-244.

Hartmann, D., Guevara, M. R., Jara-Figueroa, C., Aristarán, M., & Hidalgo, C. A. (2017a). Linking Economic Complexity, Institutions, and Income Inequality. World Development, 93, 75-93.

Hartmann, D., Jara-Figueroa, C., Guevara, M., Simoes, A., & Hidalgo, C. A. (2017b). The structural constraints of income inequality in Latin America. arXiv preprint arXiv:1701.03770.

Hausmann, R., & Hidalgo, C. A. (2011). The network structure of economic output. Journal of Economic Growth, 16(4), 309-342.

Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M., Simoes, A., & Yildirim, M. A. (2014). The atlas of economic complexity: Mapping paths to prosperity. Mit Press.

Hausmann, R., Hwang, J., & Rodrik, D. (2007). What you export matters. Journal of economic growth, 12(1), 1-25.

Hausmann, R., & Klinger, B. (2006). Structural transformation and patterns of comparative advantage in the product space.

Herfindahl, O. C. (1950). Concentration in the steel industry (Doctoral dissertation, Columbia University.).

Hidalgo, C. A., Klinger, B., Barabási, A. L., & Hausmann, R. (2007). The product space conditions the development of nations. Science, 317(5837), 482-487.

Hidalgo, C. A. (2009). The dynamics of economic complexity and the product space over a 42 year period (No. 189). Center for International Development at Harvard University.

Hidalgo, C. A., & Hausmann, R. (2009). The building blocks of economic complexity. Proceedings of the national academy of sciences, 106(26), 10570-10575.

Hidalgo, C. A. (2011). Discovering Southern and East Africa’s industrial opportunities. German Marshall Fund Economic Policy Paper Series.

Hirschman, A. O. (1945). National power and the structure of international trade.

Hirschman, A. O. (1958). The strategy of economic growth. Yale. New Haven.

Krueger, A. O. (1978). Liberalization, Direction of Bias, and Economic Growth. In Liberalization Attempts and Consequences (pp. 277-300). NBER.

Nickell, S. (1981). Biases in dynamic models with fixed effects. Econometrica: Journal of the Econometric Society, 1417-1426.

Nübler, I. (2014). A theory of capabilities for productive transformation: Learning to catch up. Transforming Economies, 113.

Ricardo, D. (1891). Principles of political economy and taxation. G. Bell and sons.

Rodriguez, F., & Rodrik, D. (2000). Trade policy and economic growth: a skeptic's guide to the cross-national evidence. NBER macroeconomics annual, 15, 261-325.

Rodrik, D. (2006). What's so special about China's exports?. China & World Economy, 14(5), 1-19.

Rodrik, D., Subramanian, A., & Trebbi, F. (2004). Institutions rule: the primacy of institutions over geography and integration in economic development. Journal of economic growth, 9(2), 131-165.

Roodman, D. (2006). How to do xtabond2: An introduction to difference and system GMM in Stata.

Roodman, D. (2009). A note on the theme of too many instruments. Oxford Bulletin of Economics and statistics, 71(1), 135-158.

Rosenstein-Rodan, P. N. (1943). Problems of industrialisation of eastern and south-eastern Europe. The economic journal, 53(210/211), 202-211.

Sachs, J. D., & Warner, A. M. (1995). Economic convergence and economic policies (No. w5039). National Bureau of Economic Research.

Saviotti, P. P., & Frenken, K. (2008). Export variety and the economic performance of countries. Journal of Evolutionary Economics, 18(2), 201-218.

Simoes, A. J. G., & Hidalgo, C. A. (2011, August). The Economic Complexity Observatory: An Analytical Tool for Understanding the Dynamics of Economic Development. In Scalable Integration of Analytics and Visualization.

Taylor, A. M. (1998). On the costs of inward-looking development: price distortions, growth, and divergence in Latin America. The Journal of Economic History, 58(1), 1-28.

Wacziarg, R. (2001). Measuring the dynamic gains from trade. The world bank economic review, 15(3), 393-429.

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. Journal of econometrics, 126(1), 25-51.

Winters, L. A. (2004). Trade liberalisation and economic performance: an overview. The Economic Journal, 114(493).

World Bank. (2016, October 19). Worldwide Governance Indicators. Retrieved from http://data.worldbank.org/data-catalog/worldwide-governance-indicators.

World Bank. (2017, April 1). World Development Indicators. Retrieved from http://data.worldbank.org/data-catalog/world-development-indicators.

Appendix

Table A1. Detailed Variable Descriptions

Code	**Label**	**Extended Notes**	**Sourced Date**	**Source**
isoa3	ISO Alpha-3	Standardized ISO Identifier	21/06/2017	ISO
iso	ISO Numeric	Standardized Numerical ISO Identifier	21/06/2017	ISO
eci	Economic Complexity Index	Economic complexity is measured by the diversity of the products that countries are able to make and the ubiquity of those products across the countries that make them.	22/05/2017	The Observatory of Economic Complexity (OEC)
im	Imports of goods and services (constant 2010 US$)	Represents the value of all goods and other market services received from the rest of the world.	01/06/2017	World Development Indicators (WDI)
ex	Exports of goods and services (constant 2010 US$)	Represents the value of all goods and other market services provided to the rest of the world.	01/06/2017	World Development Indicators (WDI)
tawm	Tariff rate, applied, weighted mean, all products (%)	Weighted mean applied tariff is the average of effectively applied rates weighted by the product import shares corresponding to each partner country.	01/06/2017	World Development Indicators (WDI)
sss	Services, etc., value added (% of GDP) [Service Sector Size]	Measures value added in net output services after adding up all outputs and subtracting intermediate inputs. Services correspond to ISIC divisions 50-99.	01/06/2017	World Development Indicators (WDI)
gdppc	GDP per capita (constant 2010 US$)	GDP per capita is gross domestic product divided by midyear population. Data are in constant 2010 U.S. dollars.	18/07/2017	World Development Indicators (WDI)
tpop	Population, total	Total population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship. Values are midyear estimates.	01/06/2017	World Development Indicators (WDI)
fdio	Foreign direct investment, net outflows (BoP, in billions of current US$)	FDI criterion is based on the ownership of 10 percent or more of voting stock. This series shows net outflows of investment from the reporting economy to the rest of the world. Data are in current U.S. dollars.	01/06/2017	World Development Indicators (WDI)

fdii	Foreign direct investment, net inflows (BoP, in billions of current US$)	FDI criterion is based on the ownership of 10 percent or more of voting stock. This series shows net inflows of investment (new investment less disinvestment) from the reporting economy to the rest of the world. Data are in current U.S. dollars.	01/06/2017	World Development Indicators (WDI)
va	Voice and Accountability (Values range between [-2.5, 2.5])	Reflects perceptions of the extent to which a country's citizens are able to participate in selecting their government, as well as freedom of expression, association, and a free media.	01/06/2017	World Governance Indicators (WGI)
psnv	Political Stability No Violence (Values range between [-2.5, 2.5])	Political Stability and Absence of Violence/Terrorism measures perceptions of the likelihood of political instability and/or politically-motivated violence, including terrorism.	01/06/2017	World Governance Indicators (WGI)
ge	Government Effectiveness (Values range between [-2.5, 2.5])	Reflects perceptions of the quality of public services, civil service, independence from political pressures, quality of policy formulation and implementation, and government credibility.	01/06/2017	World Governance Indicators (WGI)
rq	Regulatory Quality (Values range between [-2.5, 2.5])	Reflects perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development.	01/06/2017	World Governance Indicators (WGI)
coc	Control of Corruption (Values range between [-2.5, 2.5])	Reflects perceptions of the extent to which public power is used for private gain, including both petty and grand forms of corruption, and the "capture" of the state by elites and private interests.	01/06/2017	World Governance Indicators (WGI)
rol	Rule of Law (Values range between [-2.5, 2.5])	Reflects perceptions of the extent to which agents follow the rules of society, particularly the quality of contract enforcement, property rights, police, courts, and the likelihood of crime and violence.	01/06/2017	World Governance Indicators (WGI)

Table A2: Summary Statistics and Transformations

Code	Variable	Obs.	Min.	Max.	Median	Mean	Std. Dev.	Sample Period
year	Year	12,035	1960	2015				
iso	ISO Numeric Country Identifier	12,084	4	894				
eci	Economic Complexity Index	5,881	-7.2488	2.7188	-0.0441	-0.0014	1.0601	1964-2015
ln_ex	ln(Exports)	6,035	16.8315	28.4245	22.8839	22.9643	2.1556	1960-2015
ln_im	ln(Imports)	6,045	17.1870	28.6659	22.9583	23.0825	1.9554	1960-2015
tawm	Tariff rates, Applied, Weighted Means	2,972	0.0000	254.5800	4.8650	6.9671	7.9078	1988-2015
ln_gdppc	ln(GDP per Capita)	8,707	4.7518	11.8793	8.1440	8.2273	1.5285	1960-2015
ln_tpop	ln(Total Population)	11,909	8.3615	21.0390	15.2238	14.7523	2.4429	1960-2015
sss	Service Sector Size	6,436	2.4284	104.3466	52.3207	52.7265	15.6031	1960-2015
fdii	FDI inflows in billions	7,333	-29.6794	734.0103	0.1092	4.4002	22.9623	1960-2015
fdio	FDI outflows in billions	5,318	-119.3751	596.5143	0.0245	6.0169	29.8521	1960-2015
coc	Control of Corruption	3,437	-2.0575	2.5856	-0.2431	-0.0017	1.0040	1996-2015
ge	Government Effectiveness	3,433	-2.4871	2.4313	-0.1543	-0.0016	1.0031	1996-2015
psnv	Political Stability No Violence	3,444	-3.3239	1.9328	0.0966	-0.0140	0.9939	1996-2015
rol	Rule of Law	3,496	-2.6690	2.1205	-0.1497	-0.0070	1.0011	1996-2015
rq	Regulatory Quality	3,435	-2.6754	2.2629	-0.1155	-0.0058	1.0027	1996-2015
va	Voice and Accountability	3,481	-2.2862	1.8264	0.0299	-0.0012	1.0032	1996-2015

Note: all values rounded to four decimal places where applicable.

Table A3: Country List

ISO	Country	Region	ISO	Country	Region	ISO	Country	Region
8	Albania	ECA	340	Honduras	LAC	586	Pakistan	SA
12	Algeria	MENA	344	Hong Kong	EAP	591	Panama	LAC
32	Argentina	LAC	348	Hungary	ECA	600	Paraguay	LAC
36	Australia	EAP	356	India	SA	604	Peru	LAC
40	Austria	ECA	360	Indonesia	EAP	608	Philippines	EAP
68	Bolivia	LAC	372	Ireland	ECA	616	Poland	ECA
76	Brazil	LAC	376	Israel	MENA	620	Portugal	ECA
100	Bulgaria	ECA	380	Italy	ECA	634	Qatar	MENA
116	Cambodia	EAP	384	Côte d'Ivoire	SSA	642	Romania	ECA
120	Cameroon	SSA	392	Japan	EAP	682	Saudi Arabia	MENA
124	Canada	NA	400	Jordan	MENA	702	Singapore	EAP
144	Sri Lanka	SA	404	Kenya	SSA	704	Viet Nam	EAP
152	Chile	LAC	410	South Korea	EAP	710	South Africa	SSA
156	China	EAP	418	Laos	EAP	724	Spain	ECA
170	Colombia	LAC	422	Lebanon	MENA	729	The Sudan	SSA
178	The Congo	SSA	450	Madagascar	SSA	752	Sweden	ECA
188	Costa Rica	LAC	458	Malaysia	EAP	756	Switzerland	ECA
208	Denmark	ECA	480	Mauritius	SSA	764	Thailand	EAP
214	Dom. Rep.	LAC	484	Mexico	LAC	780	Trin. & Tob.	LAC
218	Ecuador	LAC	496	Mongolia	EAP	788	Tunisia	MENA
222	El Salvador	LAC	504	Morocco	MENA	792	Turkey	ECA
231	Ethiopia	SSA	508	Mozambique	SSA	818	Egypt	MENA
246	Finland	ECA	512	Oman	MENA	826	U.K.	ECA
250	France	ECA	528	Netherlands	ECA	834	Tanzania	SSA
266	Gabon	SSA	554	New Zealand	EAP	840	United States	NA
288	Ghana	SSA	558	Nicaragua	LAC	858	Uruguay	LAC
300	Greece	ECA	566	Nigeria	SSA	894	Zambia	SSA
320	Guatemala	LAC	578	Norway	ECA			

Where the regional acronyms are EAP for East Asia & Pacific (15 countries), ECA for Europe & Central Asia (20 countries), LAC for Latin America & Caribbean (18 countries), MENA for Middle East & North Africa (10 countries), NA for North America (2 countries), SA for South Asia (3 countries), and SSA for Sub-Saharan Africa (15 countries).

Table A4: Correlation coefficients between trade measures and ECI

Ln (Imports)	0.7848
Ln (Exports)	0.7526
Weighted Mean Applied Tariffs	-0.2874
Ln (GDP per capita)	0.7814
Ln (Total Population)	0.1106
Service Sector Size	0.6720
FDI Inflows	0.3186
FDI Outflows	0.3673
Control of Corruption	0.7390
Government Effectiveness	0.8213
Political Stability	0.5517
Rule of Law	0.7735
Regulatory Quality	0.7960
Voice and Accountability	0.6838

All values are rounded to four decimal places.

Table A5: Correlation coefficients between trade measures

	Weighted Mean Applied Tariffs
Ln (Imports)	- 0.2487
Ln (Exports)	-0.2385

All values are rounded to four decimal places.

Table A9 – Hausman Tests (Fixed Effects vs. Random Effects)
Trade Value [*ln (Imports) & ln (Exports)*]

H_0: difference in coefficients is not systematic vs. **H_1:** difference in coefficients is systematic

	FE vs. RE model I	**FE vs. RE model II**	**FE vs. RE model III**	**FE vs. RE model IV**
$\chi^2(k)$	101.74	114.30	97.49	74.88
Prob. > χ	0.0000	0.0000	0.000	0.000
Degrees of freedom (k)	2	5	7	13

Note: Standard Errors are unadjusted – Hausman Test cannot be used with clustered Standard Errors. All figures rounded to four decimal places. Under H_0, both estimators are consistent, albeit RE coefficients estimates are more efficient. Under the H_1, FE estimates are consistent, whereas RE estimates are not.