Accounting for Revealed Comparative Advantage: Economic Complexity Redux*

Ricardo Hausmann
Harvard Kennedy School
Mingzhi (Jimmy) Xu
INSE, Peking University
the National Bureau of Economic Research
and Growth Lab at Harvard Kennedy School

July 8, 2019

Abstract

Using the bipartite international trade networks to characterize the economic complexity has generated new insights into the patterns of economic growth across countries. The Economic Complexity Index (ECI) as charted by Hidalgo and Hausmann (2009) has been found particularly successful in explaining the difference in economic growth across countries. As staggering as its explanation power is, little has been known about the underneath factors influencing ECI. In this paper, we develop an analytically tractable framework that marries a full account of supply and demand factors to the global production network with a nested CES preference structure. The model provides formula offering a simple and intuitive decomposition splitting Balassa (1965)'s Revealed Comparative Advantage (RCA) into contributions of a number of different micro mechanisms. Based upon the decomposition, we study the relative importance of each margin embedded in ECI to explain the differences in income and inequality across countries. Our empirical results reveal that the positive correlation between complexity and income growth is mostly driven by the factors on the supply-side. By contrast, the negative association between economic complexity and inequality stems mainly from the demand factors underlying comparative advantage. We then propose a model-based instrumental variable to infer causal associations between economic complexity and per capita GDP growth. Our IV results show that a one standard deviation increase in ECI leads average income to grow faster by 0.65% to 0.72% annually.

Key Words: Comparative advantage, global trade pattern, economic complexity, price index, economic development

JEL Classification: F14, F43, O11, O4, O57

*We thank Robert Feenstra and Stephen Redding for early suggestions and help on this project. We also thank Andrew Bernard, Marc Muendler, Frank Neffke, Jose de Sousa, Anthony Venables and seminar participants at the World Bank, ATW, SED, Tsinghua University and Harvard Kennedy School. Financial support from the NBER is gratefully acknowledged, Please direct correspondence to: Mingzhi (Jimmy) Xu, Peking University and NBER, email: mingzhixu@nsd.pku.edu.cn.

1 Introduction

Research on why some countries are better in producing some products than others has been constantly pursued by economists, as its answer could reveal some fundamental source for economic growth (Hausmann and Hidalgo (2011), Lin (2012), and Sutton and Trefler (2016)). To approach this question, Balassa (1965) Revealed Comparative Advantage (RCA) index has been employed in countless explorations as a measure of the relative ability of a country to produce a good, because of being simple and intuitive. However, to understand its underlying channels is not immediately obvious, which usually requires functional form changes, and uses general equilibrium macro models based on assumptions that sometimes can be at odds (Redding and Weinstein (2017)). In contrast, we propose a unified framework for understanding the relative importance of different micro-mechanisms for Balassa (1965) RCA, for accounting for the global trade pattern, and for studying how each channel contributes to economic development, all of which have remained elusive in the existing literature.[1]

Our first main insight is to use a well-established trade nesting structure to develop a theory-consistent measure of Balassa RCA that summarizes the cost of sourcing goods from and the demand by each country and industry. All the input we need to implement this framework is the assumed demand system, its parameters, and data on unit values and expenditures, without making any of the standard modeling assumptions on supply-side.[2] To overcome the disadvantage of being mathematically non-separable for Balassa RCA, we take advantage of the similarity between arithmetic and geometric averages. Then, with our model, we show that the traditional Balassa RCA can be expressed as the contributions of a number of different micro-mechanisms in a sequence of steps:[3]

$$\ln RCA_{igt} \approx \overbrace{\underbrace{\ln\left(RCA^{P}_{igt}\right)}_{\text{Average prices}} + \underbrace{\ln\left(RCA^{\varphi^S}_{igt}\right)}_{\text{Average producer quality}} + \underbrace{\ln\left(RCA^{N}_{igt}\right)}_{\text{Variety}} + \underbrace{\ln\left(RCA^{S}_{igt}\right)}_{\text{Variety differentiation}}}^{Supply-side\ factors}$$

$$+ \underbrace{\overbrace{\ln\left(RCA^{\varphi^D}_{igt}\right)}^{\text{Average taste}} + \overbrace{\ln\left[\frac{N^{M}_{igt}/N^{EM}_{gt}}{N^{MG}_{it}/N^{EMG}_{t}}\right]}^{\text{Number of "customers"}} + \overbrace{\ln\left[RCA_{igt}\left(\mathbf{S}, \tilde{\mathbf{w}}, \mathbf{L}\right)\right]}^{\text{Average market size}}}_{Demand-side\ factors}$$

First, how good a country is in exporting in an industry depends on her average prices, product/producer quality, product varieties, and variety differentiation. Secondly, we separate out the uncovered quality into the contributions from product/producer quality and consumer taste com-

[1]Our decomposition method is similar to Redding and Weinstein (2017), who decompose a variant measure of RCA that comprises factors only on the supply-side (by construction). Differently, besides the supply-side factors as covered by Redding and Weinstein (2017), we show that demand factors, such as market size, are also essential elements for the Balassa RCA to account for the global trade pattern.

[2]For instance, the standard trade model usually assumes the productivity follows a Pareto or Fréchet distribution.

[3]Variables $\mathbf{S}$, $\tilde{\mathbf{w}}$ and $\mathbf{L}$ denote the vectors for sector shares, real income, and population of each country. The full definition of each term is provided in the theory section.

ponents, to study their relative importance in affecting consumer perceived good quality. Thirdly, in addition to supply factors, comparative advantage crucially depends on demand factors, and these channels are overlooked in a close paper by Redding and Weinstein (2017) who study a different RCA measure.

Our second contribution is to implement the decomposition method to evaluate the sources of comparative advantage and account for the observed global trade pattern systematically. We estimate quality for varieties defined as product-origin pairs, and our exercise comprehensively covers more than 190 countries for years from 2000 to 2015. Our framework enables us to distinguish between demand and supply factors and to quantify the relative importance of different micro-mechanisms that shape a country's comparative advantage, which is thus the first attempt to rationalize the observed trade flows across countries worldwide.[4] Our empirical finding highlights that both demand and supply factors are important factors to explain the world trade flows, and they jointly explain about three-quarters of total RCA variations in the sample we studied. Within supply factors, in line with Redding and Weinstein (2017), we also highlight the role of imperfect substitutes in accounting for comparative advantage, as reflected by the dominant contribution from product quality to the production advantage. Therefore, a country's comparative advantage cannot be inferred solely from conventional measures of average prices. Instead, it also depends on the other non-conventional forces such as the number of varieties, goods differentiation and market access. It exhibits a similar pattern when we evaluate the changes in global trade patter over time.

Thirdly, we innovatively apply our decomposition method to the field of economic complexity (Hidalgo and Hausmann (2009); Hausmann and Hidalgo (2011)), where we study the sources of comparative advantage that are responsible for an economy's industry diversification and the overall economic sophistication. Briefly, we construct the economic complexity index (ECI) based upon each RCA component as showing up in the decomposition formula, and the corresponding ECI will be the implied complexity index solely driven by the advantage associated with that specific mechanism. Leveraging regression analysis, we find that export advantage arising from both supply and demand channels are essential for a country's complexity. Within supply components, producer quality and product diversification within sectors are more effective to change a country's overall ECI. To our knowledge, this paper is the first to propose a structural method for exploring the determinants of economic complexity, which naturally bridges the trade literature on price index and the field of economic complexity.

We further investigate the channels via which economic sophistication correlates with income growth positively (Hidalgo and Hausmann (2009); Hausmann and Hidalgo (2011)) and inequality negatively (Hartmann, Guevara, Jara-Figueroa, Aristarán, and Hidalgo (2017)). The empirical results reveal that the positive correlation between complexity and income growth is associated with lowering marginal costs, improving product quality, and expanding goods variety within sectors,

[4] Redding and Weinstein (2017) also apply their constructed RCA to accounting for the trade pattern. Different from our paper, they only include two countries (the U.S. and Chile) in the study, and their primary goal is to study how adjustment at the firm level would affect the aggregate price index.

while not much relevant for product differentiation nor demand mechanisms. In addition, for the first time, by leveraging the auxiliary equations implied by model, we construct an instrumental variable to infer the causal impact of economic complexity on economic growth. The instrumental variable relies on the RCA variations solely explained by supply factors. Based our estimates, a one standard deviation increase in overall ECI leads the annual income growth rate to rise by 0.65% to 0.72%. By contrast, the negative association between economic complexity and inequality stems mostly from the demand factors underlying comparative advantage. The application of our approach to economic complexity provides new aspects that help understand the source of economic development.

2 Related Literature

Our paper relates to several strands of existing research. First, we contribute to the literature that aims to understand the global pattern of revealed comparative advantage. Since Balassa (1965), RCA indexes have been employed in a countless application as a measure of a country's capability of producing a good relative to others. However, due to its macro aggregation and mathematical non-separability, previous studies have been focusing on exploring various alternative measures of RCA and investigating their underlying patterns.[5] Important subsequent trials aiming to develop a measure delivering the theory-consistent property include Bowen (1983), Ballance, Forstner, and Murray (1987), Hoen and Oosterhaven (2006), Yu, Cai, and Leung (2009), and there are many more.

Not until recently, a notable strand of such attempts starts to approach the subject by relying on a quantitative Ricardian trade model to determine the appropriate RCA measures, rather than appealing to certain numerical properties of particular indexes. The representative works include Costinot, Donaldson, and Komunjer (2011), Levchenko and Zhang (2016), and French (2017), who primarily focus on different measures of RCA rather than the Balassa version to avoid the undesirable numerical properties, and are interested in technology change as the fundamental impact on comparative advantage evolution. By contrast, this paper studies the Balassa RCA, and provides a new decomposition emphasizing the different micro-mechanisms at play that are highlighted in existing trade theories. Our decomposition method is similar to Redding and Weinstein (2017), who decompose a variant measure of RCA that comprises factors only on the supply-side (by construction). Differently, besides the supply-side factors as covered by Redding and Weinstein (2017), we show that demand factors (market access), such as market size, are also important elements for the Balassa RCA to account for the global trade pattern. Our paper is unique in its focus on the theoretically-founded micro-mechanisms underlying Balassa RCA indexes, which can be employed in a wide of applications where B-RCA have been traditionally used.[6]

[5] See Yeats (1985) and 0Vollrath (1991) for an early summary and discussion on Balassa's RCA index. Studies investigating the formation of the comparative advantage from a network perspective include Hausmann and Klinger (2006) and Bahar, Hausmann, and Hidalgo (2014).

[6] For instance, interesting applications include a comprehensive evaluation of the sources of export growth across coun-

Second, our research connects with trade literature on price index and the empirical estimation for elasticities of substitution between varieties. Using the price index derived from the consumer theory to study the gains from variety was not popular until the pioneered work of Feenstra (1994), who also provides a powerful empirical method for many studies.[7] As shown in Feenstra (1994) and many other extensions, with CES preference and estimated the elasticity of substitution, the macro price index can be decomposed into a number of different micromechanisms, which is useful to explain the trade pattern. The decomposition method employed in our paper follow the similar steps as in Redding and Weinstein (2017), in a nested CES framework.[8] To estimate elasticities, we follow the GMM procedures outlined in Hottman, Redding, and Weinstein (2016) and Feenstra, Xu, and Antoniades (2017). Moving beyond this literature, we apply these structural techniques to comprehensively evaluate the fundamental pattern of global trade network, by leveraging the decomposition of the most widely used Balassa revealed comparative advantage.

In addition, our paper contributes to the rising field of economic complexity, as pioneered by Hausmann, Hwang, and Rodrik (2007); Hidalgo and Hausmann (2009). Based on the science of complexity and leveraging bipartite international trade networks, the Economic Complexity Index (ECI) is developed to capture the ability of a country to transform its production structure toward the supply of more sophisticated products. A by now well-established empirical stylized fact is that countries specializing in more sophisticated goods, i.e., with greater economic complexity, is associated with faster economic growth in the subsequent period (Rodrik (2006); Hausmann, Hwang, and Rodrik (2007); Hidalgo and Hausmann (2009); Hausmann, Hidalgo, Bustos, Coscia, Simoes, and Yildirim (2014)). Not only being predictive about future growth, ECI has also been found to correlate to inequality robustly.[9] Due to these reasons, economic complexity has emerged as a critical factor in explaining economic growth and development. However, as staggering as its explanation power is, little has been known about the underneath mechanisms/determinants of economic complexity so far. The goal of this paper aims to fill this gap. We apply our decomposition method to economic complexity to study its underlying channels, and to explore the contribution of each mechanism to various economic development.[10]

The remainder of the paper is organized as follows. In section 3, we present the theoretical framework. Section 4 describes the data for analysis, and we outline our structural estimation approach in section 5. We report and discuss our empirical results in 6, and provide conclusions in section 7.

---

tries, which has remained elusive.

[7]Subsequent important research based on Feenstra (1994) include estimating the welfare gains from trade (Broda and Weinstein (2006)), investigating the source of firm heterogeneity (Hottman, Redding, and Weinstein (2016)), detecting the procompetitive effect (Feenstra, Xu, and Antoniades (2017)), developing the micro-consistent macro price index (Redding and Weinstein (2016)),

[8]Some simplifying assumptions are adopted in this paper so that we could derive the simple and intuitive decomposition formula.

[9]Hartmann, Guevara, Jara-Figueroa, Aristarán, and Hidalgo (2017) show that countries exporting complex products tend to be more inclusive and have lower levels of income inequality than countries exporting simpler products.

[10]Notably, our theoretically-founded decomposition method can also be employed in other economic complexity index constructed in a different method, as long as RCA measures are used.

3 Theoretical Framework

An Approximation of Balassa RCA

We start by reviewing the RCA measure as developed by Balassa (1965). Through the paper, we index exporter as i , importer as j, an HS 4-digit sector as g. The revealed comparative advantage of country i in sector g and year t is denoted by RCA_{igt} that could be written as:

$$RCA_{igt} = \frac{\mathbb{M}_{igt}^{M}\left(X_{jigt}\right)/\mathbb{M}_{gt}^{EM}\left(X_{jigt}\right)}{\mathbb{M}_{it}^{MG}\left(X_{jigt}\right)/\mathbb{M}_{t}^{EMG}\left(X_{jigt}\right)} \times \frac{N_{igt}^{M}/N_{gt}^{EM}}{N_{it}^{MG}/N_{t}^{EMG}} \tag{1}$$

where $X_{jigt} > 0$ is trade flow of sector g from country i to country $j \neq i$ at time t; Ω_{igt}^{M} is the collection of importers buying goods from country i in sector g at t with $N_{igt}^{M} \equiv \left\|\Omega_{igt}^{M}\right\|$ as its measure; Ω_{gt}^{EM} is the set of trading pairs in sector g at t with measure $N_{gt}^{EM} \equiv \left\|\Omega_{gt}^{EM}\right\|$; Ω_{it}^{MG} is the set of importer-sector pairs for exporter i at t with measure $N_{it}^{MG} \equiv \left\|\Omega_{it}^{MG}\right\|$; Ω_{t}^{EMG} denotes set of all importer-exporter-sector combinations with measure $N_{t}^{EMG} \equiv \left\|\Omega_{t}^{EMG}\right\|$. Given these sets definition, we express the arithmetic mean of imports (across all foreign importers) faced by exporter i in sector g as $\mathbb{M}_{igt}^{M}\left(X_{jigt}\right) = \left(\sum_{j\in\Omega_{igt}^{M}} X_{jigt}\right)/N_{igt}^{M}$. Similarly, we can derive the arithmetic mean of the transactions (across all exporter-importer) in sector g at time t, i.e., $\mathbb{M}_{gt}^{EM}\left(X_{jigt}\right) \equiv \left(\sum\sum_{i,j\in\Omega_{gt}^{EM}} X_{jigt}\right)/N_{gt}^{EM}$, the arithmetic mean of the trade flows (across all importer-sector) faced by exporter i, i.e., $\mathbb{M}_{it}^{MG}\left(X_{jigt}\right) \equiv \left(\sum\sum_{j,g\in\Omega_{it}^{MG}} X_{jigt}\right)/N_{it}^{MG}$, and the arithmetic mean of all transactions (across all exporter-importer-sector) in year t, i.e., $\mathbb{M}_{t}^{EMG}\left(X_{jigt}\right) \equiv \left(\sum\sum\sum_{i,j,g\in\Omega_{t}^{EMG}} X_{jigt}\right)/N_{t}^{EMG}$. The original formula appearing in (1) shows that a country's RCA depends on her average intensive margin in sector g as well as his prevalence in the world market, without revealing further detailed information. Though being intuitive and straightforward, the arithmetic averaging makes it challenging to connect the original B-RCA to trade models so that a further decomposition to investigate itsunderlying channels cannot be well executed (French (2017)).

To overcome this issue and for tractability, we use the geometric average of trade flows in (1), instead of the arithmetic average, hence modifying the B-RCA to a new formula provided in (2). The approximation allows us to employ a nested CES framework similar to Redding and Weinstein (2017) so that we could decompose RCA into the contributions of different margins in a sequence of steps.

$$RCA_{igt} \approx \frac{\tilde{\mathbb{M}}_{igt}^{M}\left(X_{jigt}\right)/\tilde{\mathbb{M}}_{gt}^{EM}\left(X_{jigt}\right)}{\tilde{\mathbb{M}}_{it}^{MG}\left(X_{jigt}\right)/\tilde{\mathbb{M}}_{t}^{EMG}\left(X_{jigt}\right)} \times \frac{N_{igt}^{M}/N_{gt}^{EM}}{N_{it}^{MG}/N_{t}^{EMG}} \tag{2}$$

In the new formula (2), all the variables have the same meanings as in (1), except that we use the geometric mean to average trade flows across all foreign importers faced by exporter i in sector g, i.e., $\tilde{\mathbb{M}}^{M}_{igt}\left(X_{jigt}\right) = \left(\prod_{j\in\Omega^{M}_{igt}} X_{jigt}\right)^{1/N^{M}_{igt}}$, across all importer-sector faced by exporter i, i.e., $\tilde{\mathbb{M}}^{MG}_{it}\left(X_{jigt}\right) \equiv \left(\prod\prod_{j,g\in\Omega^{MG}_{it}} X_{jigt}\right)^{1/N^{MG}_{it}}$, across all exporter-importer in sector g, i.e., $\tilde{\mathbb{M}}^{EM}_{gt}\left(X_{jigt}\right) \equiv \left(\prod\prod_{i,j\in\Omega^{EM}_{gt}} X_{jigt}\right)^{1/N^{EM}_{gt}}$, and finally across all exporter-importer in sector g, i.e., $\tilde{\mathbb{M}}^{EMG}_{t}\left(X_{jigt}\right) \equiv \left(\prod\prod\prod_{i,j,g\in\Omega^{EMG}_{t}} X_{jigt}\right)^{1/N^{EMG}_{t}}$. Figure 1 presents the fit performance for using geometric mean in B-RCA. Each industry corresponds to a four digit code under HS classification. The horizontal axis is the approximation using geometric mean as implied by (2), and the vertical axis is Balassa (1965)'s RCA as shown in (1). The figures show that the approximation performs well and stable across years, i.e., there is a strong positive correlation between the original and approximated RCA as indicated the high R-square raning from 0.67 to 0.77.[11] The approximation performance is robust to using alternative industry classification such as four-digit SITC rev2 code, whose fitting performance is provided in appendix Figure A.1.

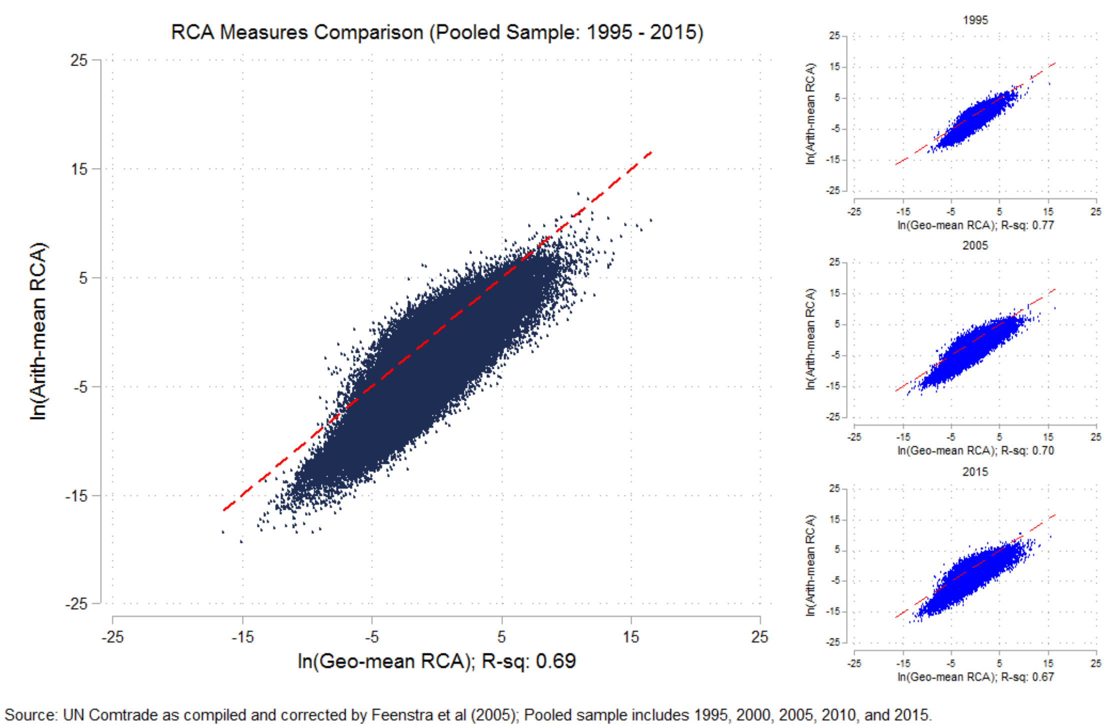


Figure 1: RCA Measures Comparison (HS 4-digit)

The approximation described above allows us to transform the B-RCA into a log-additive forma-

[11]The two measures seem to differ systematically in the intersect (i.e., level). However, it will not impose much trouble in our following analysis since one could normalize the $\ln \widehat{RCA}_{igt}$ (the approximated measure) such that its average value is the same as $\ln RCA_{igt}$ (the original measure).

tion for traceability purpose. Next, we introduce our model structure to further split the product of trade flows, as captured by $\tilde{\mathbb{M}}\left(X_{jigt}\right)$, into different variables of interest.

Demand Strucutre

We study a multi-country economy in the international framework. The preference of the representative consumer in each country is nested CES similar to Redding and Weinstein (2017). The first level CES determines the demand for each sector that consists of both tradable and non-tradable ones. Within each sector, second level CES nesting determines the demand for differentiated varieties that can be sourced from domestic or from international markets. In practice, we define a variety as country–6 digit HS code pair.

Given the preference structure, the aggregate unit expenditure for country j in year t is defined over the sectoral price index P^G_{jgt} for each sector $g \in \Omega^G$:

$$P_{jt} = \left[\sum_{g\in\Omega^G} \left(P^G_{jgt}\right)^{1-\sigma^G}\right]^{\frac{1}{1-\sigma^G}}, \sigma^G > 0 \tag{3}$$

where Ω^Gdenotes the set of all sectors that contains both tradable and non-tradable sectors., and σ^G is the elasticity of substitution across sectors. Moving forward, within each sector there are varieties of different kinds. We follow the Armington assumption that goods (six-digit HS code) produced by different countries are close but not perfect substitutes, and a variety is thus defined by a unique country -6-digit HS code pair. The unit expenditure (P^G_{jgt}) for sector g depends on the prices (P^V_{vt}) and demand parameters (φ^V_{vt}) for each variety $v \in \Omega^V_{jigt}$ and $\forall i \in \Omega^I_{jgt}$:

$$P^G_{jgt} = \left[\sum_{i\in\Omega^I_{jgt}} \sum_{v\in\Omega^V_{jigt}} \left(P^V_{vt}/\varphi^V_{vt}\right)^{1-\sigma^V_g}\right]^{\frac{1}{1-\sigma^v_g}}, \sigma^V_g > 1, \varphi^V_{vt} > 0 \tag{4}$$

where Ω^I_{jgt} denotes collection of countries (including j self) selling in country j, sector g in year t; Ω^V_{jigt} denotes the available variety set in sector g of country j that are supplied by country i; σ^V_g is the elasticity of substitution across varieties for sector g, and φ^V_{vt} captures the relative demand for each variety v.[12] In line with Redding and Weinstein (2017), we assume the unit expenditure function within each sector have the same forms for both final consumption and intermediate use so that we could aggregate both sources of expenditure. As highlighted in Redding and Weinstein (2017), one advantage of the framework is that it allows some sectors to be non-traded, and allow the existence of both domestic and foreign varieties within tradable sectors.

Without knowing transactions for non-tradable sectors or for the domestic varieties within tradable sectors, we recover them by using the expenditure share of tradable sectors (μ^T_{jt}), the expenditure share of the imported varieties within tradable sectors (μ^G_{jgt}), the aggregate price indexes of the trad-

[12]We do not impose any restriction on the demand shifter. For instance, it can contain a country-sector specific component φ^G_{jg}.

able sectors ($\mathbb{P}^T_{jt}$) and of the imported varieties within tradable sectors ($\mathbb{P}^G_{jgt}$), as well as the detailed trade data. To see this, we rewrite (3) as:

$$P_{jt} = \left(\mu^T_{jt}\right)^{\frac{1}{\sigma^G - 1}} \mathbb{P}^T_{jt}, \quad \mu^T_{jt} = \frac{\sum_{g\in\Omega^T}\left(P^G_{jgt}\right)^{1-\sigma^G}}{\sum_{g\in\Omega^G}\left(P^G_{jgt}\right)^{1-\sigma^G}}, \quad \mathbb{P}^T_{jt} = \left[\sum_{g\in\Omega^T}\left(P^G_{jgt}\right)^{1-\sigma^G}\right]^{\frac{1}{1-\sigma^G}} \tag{5}$$

and rewrite (4) as:

$$P^G_{jgt} = \left(\mu^G_{jgt}\right)^{\frac{1}{\sigma^V - 1}} \mathbb{P}^G_{jgt}, \quad \mu^G_{jgt} = \frac{\sum_{i\in\Omega^E_{jgt}}\sum_{v\in\Omega^V_{jigt}}\left(P^V_{vt}/\varphi^V_{vt}\right)^{1-\sigma^V_g}}{\sum_{i\in\Omega^I_{jgt}}\sum_{v\in\Omega^V_{jigt}}\left(P^V_{vt}/\varphi^V_{vt}\right)^{1-\sigma^V_g}}, \quad \mathbb{P}^G_{jgt} = \left[\sum_{i\in\Omega^E_{jgt}}\sum_{v\in\Omega^V_{jigt}}\left(P^V_{vt}/\varphi^V_{vt}\right)^{1-\sigma^V_g}\right]^{\frac{1}{1-\sigma^V_g}} \tag{6}$$

where $\Omega^T \subset \Omega^G \left(\mu^T_{jt}\right)$ denotes the set of tradable sectors, and $\Omega^E_{jgt} \equiv \left\{\Omega^I_{jgt} : i \neq j\right\}$ is the subset of foreign countries supplying importer j within sector g in year t. To understand the margins affecting country's export performance as revealed by RCA, in the following analysis, we first examine the contribution of individual countries to trade pattern and aggregate prices.

It is convenient to rewrite the sectoral import price index ($\mathbb{P}^G_{jgt}$) appearing in (6) in terms of price indexes of each foreign exporting country within that sector ($\mathbb{P}^E_{jigt}$):

$$\mathbb{P}^G_{jgt} = \left[\sum_{i\in\Omega^E_{jgt}}\left(\mathbb{P}^E_{jigt}\right)^{1-\sigma^V_g}\right]^{\frac{1}{1-\sigma^V_g}}, \quad \mathbb{P}^E_{jigt} = \left[\sum_{v\in\Omega^V_{jigt}}\left(P^V_{vt}/\varphi^V_{vt}\right)^{1-\sigma^V_g}\right]^{\frac{1}{1-\sigma^V_g}} \tag{7}$$

This exporter price index (7) is a key object in our empirical analysis, which summarizes exporter i's ability to sell products to importer j within sector g at time t. Using the properties of CES demand and apply it to (7), the share of variety v in the expenditure on each exporter s^V_{vt} is given by:

$$s^V_{vt} = \frac{\left(P^V_{vt}/\varphi^V_{vt}\right)^{1-\sigma^V_g}}{\sum_{v\in\Omega^V_{jigt}}\left(P^V_{vt}/\varphi^V_{vt}\right)^{1-\sigma^V_g}} \tag{8}$$

where exporter and sector expenditure shares are defined analogously. We can rearrange the share term of (8) using the country price index as(7) to obtains:

$$\mathbb{P}^E_{jigt} = \frac{P^V_{vt}}{\varphi^V_{vt}}\left(s^V_{vt}\right)^{\frac{1}{\sigma^V_g - 1}}, \quad \forall v \in \Omega^V_{jigt} \tag{9}$$

Taking logarithms on both sides, averaging across varieties within country and sector, and adding and subtracting (8), we obtain the following exact log-linear decomposition of the CES price index into four terms ($\forall v \in \Omega^V_{jigt}$):

$$\ln \mathbb{P}^{E}_{jigt} = \underbrace{\mathbb{E}^{V}_{jigt}\left[\ln P^{V}_{vt}\right]}_{\text{Average prices}} - \underbrace{\mathbb{E}^{V}_{jigt}\left[\ln \varphi^{V}_{vt}\right]}_{\text{Average quality \& Taste}} + \underbrace{\frac{1}{\sigma^{V}_{g}-1}\left(\mathbb{E}^{V}_{jigt}\left[\ln s^{V}_{vt}\right] - \ln \frac{1}{N^{V}_{jigt}}\right)}_{\text{Dispersion of quality-adjusted prices}} - \underbrace{\frac{1}{\sigma^{V}_{g}-1}\ln N^{V}_{jigt}}_{\text{Variety}}, \tag{10}$$

where $\mathbb{E}^{V}_{jigt}$ denotes arithmetic operator so that $\mathbb{E}^{V}_{jigt}\left[\ln P^{V}_{vt}\right] = \frac{1}{N^{V}_{jigt}}\sum_{v\in\Omega^{V}_{jigt}} \ln P^{V}_{vt}$; and the superscript V in $\mathbb{E}^{V}_{jigt}$ indicates the mean is taken across varieties; and subscript $jigt$ indicates the averaging operator is applied to varieties within importer j, exporter i, sector g and year t. The expression for sectoral export price index in (9) highlights the imperfect substitutes across varieties (Redding and Weinstein (2017)), which depends on the number of varieties ($\ln N^{V}_{jigt}$) and dispersion of quality-adjusted prices ($\mathbb{E}^{V}_{jigt}\left[\ln s^{V}_{vt}\right] - \ln 1/N^{V}_{jigt}$) in addition to the product price ($\mathbb{E}^{V}_{jigt}\left[\ln P^{V}_{vt}\right]$) and quality ($\mathbb{E}^{V}_{jigt}\left[\ln \varphi^{V}_{vt}\right]$).[13]

RCA Decomposition

So far, we have been focused on the price indexes that determine the costs of sourcing goods from a given exporter for a given sector. As trade flows are pinned down by relative price indexes, it is straightforward to translate the exporter price indexes into the determinants of trade patterns across countries and sectors. In the following analysis, we derive a theoretically-rigorous measure of RCA based on the CES demand system described above. For convenience, we first introduce the operator $\Xi_{igt}\left(y_{jigt}\right)$ as:

$$\Xi_{igt}\left(y_{jigt}\right) \equiv \frac{\tilde{\mathbb{M}}^{M}_{igt}\left(y_{jigt}\right)/\tilde{\mathbb{M}}^{EM}_{gt}\left(y_{jigt}\right)}{\tilde{\mathbb{M}}^{MG}_{it}\left(y_{jigt}\right)/\tilde{\mathbb{M}}^{EMG}_{t}\left(y_{jigt}\right)}$$

where $y_{jigt} > 0$ can be any variable that is i, j, g and t specific; the several operators $\tilde{\mathbb{M}}\left(y_{jigt}\right)$ follows the same definition as in section 3.

Importer j's expenditure on export i ($i \neq j$) as a share of its expenditure on all foreign exporters within sector g in year t as:

$$\mathbb{S}^{E}_{jigt} = \frac{\sum_{v\in\Omega^{V}_{jigt}}\left(P^{V}_{vt}/\varphi^{V}_{vt}\right)^{1-\sigma^{V}_{g}}}{\sum_{h\in\Omega^{E}_{jgt}}\sum_{v\in\Omega^{V}_{jhgt}}\left(P^{V}_{vt}/\varphi^{V}_{vt}\right)^{1-\sigma^{V}_{g}}} = \frac{\left(\mathbb{P}^{E}_{jigt}\right)^{1-\sigma^{V}_{g}}}{\left(\mathbb{P}^{G}_{jgt}\right)^{1-\sigma^{V}_{g}}} = \frac{X_{jigt}}{\sum_{h\in\Omega^{E}_{jgt}} X_{jigt}}, \quad i \neq j \tag{11}$$

where the numerator denotes importer j's price index for exporting country i in sector g at time t ($\mathbb{P}^{E}_{jigt}$); the denominator captures importer j's overall import price index in sector g at time t ($\mathbb{P}^{G}_{jgt}$). Then we use $\mathbb{S}^{E}_{jigt}$ and country j's total import in sector g (i.e.,$\mathbb{X}^{E}_{jgt} = \sum_{i\in\Omega^{E}_{jgt}} X_{jigt}$) to substitute for the trade flow from i to j in sector g (i.e., X_{jigt}) in the adjusted RCA_{igt} appearing in equation 2 :

[13] In case of $\sigma^{V}_{g} > 1$, a bigger number of goods sold by exporter i or greater product differentiation for exporter i will decrease the sectoral export price index.

$$
\begin{aligned}
RCA_{igt} &= \Xi_{igt}\left(X_{jigt}\right) \times \frac{N^{M}_{igt}/N^{EM}_{gt}}{N^{MG}_{it}/N^{EMG}_{t}} \\
&= \Xi_{igt}\left(\mathbb{S}^{E}_{jigt}\right) \times \Xi_{igt}\left(\mathbb{X}^{E}_{jgt}\right) \times \frac{N^{M}_{igt}/N^{EM}_{gt}}{N^{MG}_{it}/N^{EMG}_{t}}
\end{aligned} \tag{12}
$$

where N^{M}_{igt} is the number of destinations importing goods g from i; $N^{E,M}_{gt}$ is the number of country pairs with positive trade flow in sector g; $N^{M,G}_{it}$ is the number of destination-sector pairs with positive trade flow sourcing from country i; $N^{E,M,G}_{t}$ denotes the number of importer-exporter-sector pairs with positive trade flow.

As we now show, the relative terms appearing in (12) will allow us to quantify the different economic mechanisms in understanding patterns of trade across countries and sectors. Using (11) and (12), we derive the decomposition of a country's RCA as:

$$
RCA_{igt} = \Xi_{igt}\left(\left[\mathbb{P}^{E}_{jigt}\right]^{1-\sigma^{V}_{g}}\right) \times \Xi_{igt}\left(\mathbf{S}, \tilde{\mathbf{w}}, \mathbf{L}\right) \times \frac{N^{M}_{igt}/N^{EM}_{gt}}{N^{MG}_{it}/N^{EMG}_{t}} \tag{13}
$$

and

$$
\Xi_{igt}\left(\mathbf{S}, \tilde{\mathbf{w}}, \mathbf{L}\right) \equiv \Xi_{igt}\left(\left[S_{jgt}\right]^{\frac{\sigma^{V}_{g}-\sigma^{G}}{1-\sigma^{G}}}\right) \times \Xi_{igt}\left(\left[w_{jt}/P_{jt}\right]^{1-\sigma^{V}_{g}}\right) \times \Xi_{igt}\left(w_{jt}^{\sigma^{V}_{g}} L_{jt}\right) \tag{14}
$$

where $\mathbf{S}$, $\tilde{\mathbf{w}}$ and $\mathbf{L}$ denote the vector of sector shares, real income, and population. The first component in (13) denotes the relative costs to an importer of sourcing goods across countries and sectors, as reflected by the relative price index adjusted by the elasticity of substitution; the second component captures the factors related to market size that affect RCA_{igt}. Specifically, $\Xi_{igt}\left(\left[S_{jgt}\right]^{\frac{\sigma^{V}_{g}-\sigma^{G}}{1-\sigma^{G}}}\right)$ captures the average expenditure share on sector g across the global markets; $\Xi_{igt}\left(\left[w_{jt}/P_{jt}\right]^{1-\sigma^{V}_{g}}\right)$ capture the average real income across destination markets depending on the nominal average income and price level, and $\Xi_{igt}\left(w_{jt}^{\sigma^{V}_{g}} L_{jt}\right)$ reflects the total market size adjusted by the elasticity of substitution. All the variables appearing in (14), i.e., sector shares S_{jgt}, real incomes w_{jt}/P_{jt}, nominal per capita GDP w_{jt} and population L_{jt} are observables in the data. The last component in (13) captures the relative number of "customers" a country is trading with. [14]

To breakdown the price index $\mathbb{P}^{E}_{jigt}$ in (13), we first take logs on both sides of the equation (13) and subsitute $\ln\left(\mathbb{P}^{E}_{jigt}\right)$ using equation (10), which yields the exact log-linear decomposition of RCA as:

[14] In Redding and Weinstein (2017), they provide a decomposition of an RCA index consisting of supply factors only. By contrast, the B-RCA includes both demand and supply factors, which are both critical to account for the global trade pattern.

$$\ln RCA_{igt} \approx \overbrace{\underbrace{\ln\left(RCA^{P}_{igt}\right)}_{\text{Average prices}} + \underbrace{\ln\left(RCA^{\varphi^S}_{igt}\right)}_{\text{Average quality}} + \underbrace{\ln\left(RCA^{S}_{igt}\right)}_{\text{Dispersion quality-adjusted prices}} + \underbrace{\ln\left(RCA^{N}_{igt}\right)}_{\text{Variety}}}^{Supply-side\ factors}$$

$$\underbrace{+\overbrace{\ln\left(RCA^{\varphi^D}_{igt}\right)}^{\text{Average Taste}} + \overbrace{\ln\left[\frac{N^{M}_{igt}/N^{EM}_{gt}}{N^{MG}_{it}/N^{EMG}_{t}}\right]}^{\text{Number of "Customers"}} + \overbrace{\ln\left[RCA_{igt}\left(\mathbf{S}, \tilde{\mathbf{w}}, \mathbf{L}\right)\right]}^{\text{Average Market Size}}}_{Demand-side\ factors} \tag{15}$$

where each margin is defined in Appendix B. The decomposition in (15) provides a close link between theory and data. The formula offers intuitive and straightforward decomposition splitting Balassa (1965)'s RCA into contributions of a number of different micro-mechanisms. Using the decomposition equation, we are now able to study the relative importance of each margin in explaining the global trade pattern. Further, we apply our approach to the field of economic complexity, investigating the corresponding margins that explain cross-country differences in economic growth and inequality.

4 Data Description

To undertake our empirical analysis exploring the determinants of global trade patterns and economic complexity, we use international trade flow data that are readily available from the United Nations Comtrade Database for years between 2000 and 2015. We include a wide coverage of countries and HS 4-digit industries in our analysis, i.e., there are 165 countries and 1244 industries in the year 2000, and 154 countries and 1240 industries in the year 2015. In the data, we follow the Armington assumption by considering an exporter-HS 6-digit combination as a unique variety. For the years that we have world trade data, we also use data on unit prices of varieties (in 6-digit HS code) from the Trade Unit Values Database maintained by CEPII.[15] The Trade Unit Values database contains unit value information (measured in US dollars per ton) with a maximum of 173 reporters, 255 partners, and more than 5,000 product categories per year. We use the Cost of Insurance and Freight (CIF) unit values that rely on importers' declarations and include all trade costs (except tariffs and domestic taxes after the border).

Besides trade data, information on sectoral production across countries is from the National Accounts Main Aggregates Database. Though the sector is more aggregate, compared to an industry defined as an HS 4-digit code, the main database provides us with the most comprehensive sector coverage of 216 countries and over 38 years. [16] Other country variables, such as real income, pop-

[15]Details of Trade Unit Database refer to Berthou and Emlinger (2011). The database is also used in various recent studies such as Blonigen (2015) and Chen and Juvenal (2016).

[16]The alternative sources for sectoral output, such as Industrial Demand-Supply Balance Database from UNIDO, surfer the disadvantage of missing values for many industries and countries.

ulation, and GDP, are from Worldwide Governance Indicators, World Development Indicators, and the Penn World Table.

5 Structural Estimation

In order to apply our approach to data, we need estimate the elasticities of substitution (σ^G, σ_g^V), and recover consumer perceived quality for varieties (φ_{vt}^V). Similar to Redding and Weinstein (2017), we now require some structures about the supply-side. The detailed procedures will be discussed subsequently in greater depth.

5.1 Lowest-tier of Demand

Estimation of σ_g^V in the lowest-tier of demand follows the standard approach proposed by Feenstra (1994), which are later improved and applied by various studies such as Broda and Weinstein (2006); Hottman, Redding, and Weinstein (2016); Feenstra, Xu, and Antoniades (2017); Redding and Weinstein (2017). The identification is as follows. The slopes of demand and supply curves for a given industry are assumed to be constant across varieties and over time, but their intercepts are allowed to vary across varieties and time. If the supply and demand intercepts for a given variety are orthogonal, there exists a rectangular hyperbola in demand elasticity and supply elasticity space[17] that could best fit the observed price and sales of that variety.[18] We will describe the orthogonality conditions for each variety in terms of double-differenced supply and demand intercepts in the rest of the section, and outline the GMM procedures used for estimating the demand and supply elasticities for each industry.

Starting from equation (8) that characterizes the share of variety v in the expenditure on each exporter, we take the time difference and the difference relative to another variety (i.e., reference variety) consumed in the same industry and by the same importer. The double-differencing yields:

$$\Delta^{k,t} \ln s_{vt}^V = \left(1 - \sigma_g^V\right) \Delta^{k,t} \ln p_{vt}^V + \omega_{vt} \tag{16}$$

where k indicates the reference variety consumed by the same importer in industry g. The unobserved error term $\omega_{vt} \equiv \left(1 - \sigma_g^V\right) \left[\Delta^t \ln \varphi_{kt}^V - \Delta^t \ln \varphi_{vt}^V\right]$ captures the idiosyncratic double differenced demand shocks. Next, we assume the total variable costs for exporter i to supply importer j with variety v is

$$V_{jivt}\left(Q_{jivt}\right) = z_{jivt} Q_{jivt}^{1+\delta_g} \tag{17}$$

[17]Supply elasticity is the elasticity of variable cost with respect to quantity which is introduced in equation (17).

[18]The orthogonality assumption alone does not allow identification, as there is always a tradeoff between the value of demand elasticity and supply elasticity. But if the variance of demand and supply intercepts are heteroskedastic across varieties in the industry, the hyperbola best fitting the data will differ across varieties. Since the slopes of the demand and supply curves are the same, the intersection of the hyperbolas of different varieties in the industry separately identifies the demand and supply elasticities.

where Q_{jivt} denotes the total quantity of variety v supplied by exporter i in importing country j; parameter δ_g determines the convexity of marginal cost with respect to output for varieties in industry g; z_{jivt} is an importer-exporter-variety specific shifter of the cost function. Costs are paid in terms of a composite factor input that is chosen the numeraire. We require δ_g to be greater than zero, which will be estimated. We consider the global market structure is perfect competition[19] so that price equals marginal cost,[20] i.e., $p^V_{vt} = \tau_{jigt}\left(1+\delta_g\right) z_{vt} Q^{\delta_g}_{vt}$. Given that $Q_{vt} = S^V_{vt} X_{jigt} / p^V_{vt}$ where X_{jigt} denotes the total purchase by importer j from exporter i in sector g, the pricing equation can be written in double-differenced form as:

$$\Delta^{k,t} \ln P^V_{vt} = \frac{\delta_g}{1+\delta_g} \Delta^{k,t} \ln s^V_{vt} + \kappa_{vt} \tag{18}$$

where the unobserved error term $\kappa_{vt} \equiv \frac{1}{1+\delta_g}\left[\Delta^t \ln z_{vt} - \Delta^t \ln z_{kt}\right]$ captures the idiosyncratic double differenced supply shocks. The orthogonality condition for each variety is then defineds as:

$$G\left(\boldsymbol{\beta}_g\right) = \mathbb{E}_T\left[\omega_{vt}\left(\boldsymbol{\beta}_g\right)\kappa_{vt}\left(\boldsymbol{\beta}_g\right)\right] = 0 \tag{19}$$

where $\boldsymbol{\beta}_g = \begin{pmatrix} \sigma^V_g \\ \delta_g \end{pmatrix}$. The condition provided by (19) assumes that the idiosyncratic demand and supply shocks at the variety level are independent, since the variety and variety-year fixed effects have been differenced out. For each industry g, we stack the orthogonality conditions to form the GMM objective function:

$$\hat{\boldsymbol{\beta}}_g \quad = \quad argmax_{\boldsymbol{\beta}_g}\left\{G^*\left(\boldsymbol{\beta}_g\right)' W G^*\left(\boldsymbol{\beta}_g\right)\right\} \tag{20}$$

where $G^*\left(\boldsymbol{\beta}_g\right)$ denotes counterpart of $G\left(\boldsymbol{\beta}_g\right)$ in the data that are stacked over all varieties in industry g, and W is a positive defnite weighting matrix. Following Feenstra, Xu, and Antoniades (2017), we give more weight to varieties which has bigger quantity in the data.

After we obtain the elasticities, we then recover the taste parameters. The share of variety v in the expenditure on all exporter s^V_{vt} is given by:

$$s_{jivt} = \frac{\left(P^V_{jivt}/\varphi^V_{jivt}\right)^{1-\sigma^V_g}}{\sum_{i\in\Omega^E_{jgt}}\sum_{v\in\Omega^V_{jigt}}\left(P^V_{jivt}/\varphi^V_{jivt}\right)^{1-\sigma^V_g}} \tag{21}$$

We express the taste parameters relative to the variety whose sales is at the median level, and we recover the taste parameters from the relationship characterized below:

[19] One can impose a different market structure so that price will also depend on some markup, which is an important and interesting extension.

[20] We omit the subscript i and j for simplicity.

$$\ln\left(s_{jivt}/s_{jvt}^{med}\right) = \left(1-\sigma_g^V\right)\ln\left[P_{jivt}^V/P_{jvt}^{V,med}\right] + \left(\sigma_g^V - 1\right)\ln\left(\varphi_{jivt}^V/\varphi_{jvt}^{V,med}\right) \tag{22}$$

where s_{jvt}^{med} and $P_{jvt}^{V,med}$ denote the sales share and price for the variety whose sales is at the median level in industry g and country j. We recover variety quality using (22) and normalize quality relative to that of the median-sales variety, i.e, $\ln\varphi_{jvt}^{V,med} = 0$ by normalization. We further split variety v's consumer perceived quality into two components as provided below:[21]

$$\ln\left(\varphi_{jivt}^V\right) = \ln\left(\varphi_{jvt}^V\right) + \ln\left(\varphi_{ivt}^V\right) + \epsilon_{jivt} \tag{23}$$

where $\ln\left(\varphi_{jvt}^V\right)$ denotes the importer-specific component in quality which we interprete as *consumer taste*, and $\ln\left(\varphi_{ivt}^V\right)$ stands for the exporter-specific component in quality which we interprete as *product quality*. The residual term ϵ_{jivt} denotes any other idiosyncratic reasons affecting importer i 's perceived quality of variety v produced by country j. Computationally, importer-specific component $\ln\left(\varphi_{jvt}^V\right)$ is caculated by averaging $\ln\left(\varphi_{jivt}^V\right)$ across all varieties purchased by country j in sector g; exporter-specific component $\ln\left(\varphi_{ivt}^V\right)$ is caculated by averaging $\ln\left(\varphi_{jivt}^V\right)$ across all varieties sold by country i in sector g; the idiosyncratic taste shocks ϵ_{jivt} is what's left after substracting the importer and exporter-specific components from the recovered variety quality.

5.2 Upper-tier of Demand

With the estimates of σ_g^V and taste parameters, we could calculate the import price index $\mathbb{P}_{jgt}^G = \left[\sum_{i\in\Omega_{jgt}^E}\sum_{v\in\Omega_{jigt}^V}\left(P_{vt}^V/\varphi_{vt}^V\right)^{1-\sigma_g^V}\right]^{\frac{1}{1-\sigma_g^V}}$, and we now turn to estimate the elasticities across industries σ^G. Recall that the expenditure share of industry g in country j as $S_{jgt} = \left(P_{jgt}^G\right)^{1-\sigma^G}/\left(P_{jt}\right)^{1-\sigma^G}$, and the import share of industry g as $\mu_{jgt}^G = \left(\mathbb{P}_{jgt}^G\right)^{1-\sigma^V}/\left(P_{jgt}^G\right)^{1-\sigma^V}$, one could rewrite the industry share as:

$$S_{jgt} = \left(\mu_{jgt}^G\right)^{\frac{1-\sigma^G}{1-\sigma_g^V}}\left(\mathbb{P}_{jgt}^G\right)^{1-\sigma^G}\left(P_{jt}\right)^{\sigma^G-1} \tag{24}$$

We take the time difference and the difference relative to another industry (let it be h) in country j. The double-differencing of (24) yields:

$$\Delta^{h,t}\ln S_{jgt} = \left(\frac{1-\sigma^G}{1-\sigma_g^V}\right)\Delta^{h,t}\ln\mu_{jgt}^G + \left(1-\sigma^G\right)\Delta^{h,t}\ln\mathbb{P}_{jgt}^G \tag{25}$$

Instead of directly estimating σ^G by running an OLS on (25), one can also use an instrumental variables approach. Noting that import price index $\mathbb{P}_{jgt}^G$ can be written as:

[21] The full expression of variety v's consumer perceived quality is $\ln\left(\varphi_{jivt}^V\right) - \ln\left(\varphi_{jvt}^{V,med}\right)$, and we use $\ln\left(\varphi_{jivt}^V\right)$ for notation simplicity.

$$\Delta^{h,t} \ln \mathbb{P}^{G}_{jgt} = \underbrace{\left(\frac{1}{1-\sigma^{V}_{g}}\right) \Delta^{h,t} \ln \overline{s_{vt}} + \Delta^{h,t} \ln \overline{P^{V}_{vt}}}_{\text{Instrument}} - \Delta^{h,t} \ln \overline{\varphi^{V}_{vt}}$$

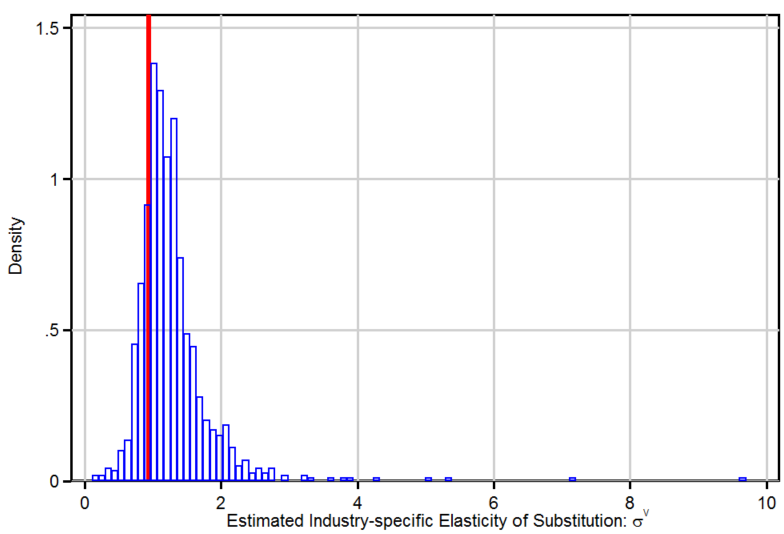
where $\overline{s_{vt}}$ and $\overline{P^{V}_{vt}}$ stand for the geometric average of variety shares and prices across all varieties within the import basket of importer j in industry g. We use the first two terms as the instrument to $\Delta^{h,t} \ln \mathbb{P}^{G}_{jgt}$. To apply the decomposition method, we require two share terms that are not avaible in data, namely, country j's expenditure share for industry g (HS 4-digit) S_{jgt}, and import share in industry g (HS 4-digit) μ^{G}_{jgt}. Alternatively, we infer these statistiscs by combined using United Nations Comtrade Database (UN Comtrade), and the National Accounts Main Aggregate Database (NAMAD). The details are provided in Appendix C.

6 Empirical Results

6.1 Estimation Results

Elasticities of Substitution

In Table 1 and Figure 2, we report our baseline estimates of the two elasticities of substitution. Since we estimate variety elasticity for each HS 4-digit sector (there are 1244 in total), it would needlessly clutter the paper to report all the details individually. Therefore we present the distribution of variety elasticities in Figure 2 and report quantiles of the distributions of elasticities (σ^{V}_{g}) across sectors and the single estimated elasticity of substitution across sectors (σ^{G}).



Note: The red line displays the elasticity of substitution across industries which equals 0.94.

Figure 2: Estimated Elasticities of Substitution

The estimated sector elasticity is 0.94 which is depicted as the red line in Figure 2. On average, the estimated elasticities across varieties are greater than the sector elasticity, with mean and median values as 1.28 and 1.18. In addition, according to Table 1, the maximum and minimum values for variety elasticities are 0.11 and 9.71, respectively. Though we do not impose any restrictions on the estimation, we find a natural ordering on average, where varieties are more substitutable within industries than across industries.

Our estimates are similar and consistent to various previous work such as Erkel-Rousse and Mirza (2002),Tokarick (2010), and Bussière, Callegari, Ghironi, Sestieri, and Yamano (2013). [22] These estimates imply that varieties within sectors are imperfect substitutes for one another, which will have important implication for accounting for the comparative advantage for a country and sector, because the observed variety unit costs along are no longer sufficient statistics for the cost of sourcing goods across countries and sectors.

Table 1: Estimated Elasticities of Substitution (World Trade Flows, 2000-2015)

Percentile	Elasticity Across Variety (σ_g^V)
Mean	1.28
Min	0.11
25th	1.00
50th	1.18
75th	1.42
Max	9.71

Note: Estimated elasticities of substitution across varieties are from GMM procedure discussed in section 5. Sectors are 4-digit HS codes; Varieties are exporter-HS 6 digit codes within importers' sectors.

Recovered Product Quality

With the estimated elasticities, we recover the variety quality and further split it into "consumer taste", "producer quality", and the unexplained component, as discussed in section 5. Table 2 summarizes the contributions of each component to account for the recovered variety quality. In the first row, we report the partial correlation coefficient that measures the association between each component and overall variety quality after removing the confounding factor between the two.[23] According to the statistics, the producer-specific component is more correlated to the overall variety quality than that carried by consumer taste, as demonstrated by the bigger partial correlation coefficient for producer quality (0.522) than that of consumer taste (0.194).

[22] Our variety elasticity of substitution (σ_g^V) also equals the import elasticity, which is estimated to vary between 0.750 and 1.255 implied by the fixed effect models in Erkel-Rousse and Mirza (2002). Note that our estimates are smaller in magnitude than those for SITC 4-digit industries as estimated by Feenstra and Romalis (2014), who assume a non-homothetic demand system and apply multiple filtering procedures to refine the data for estimation.

[23] For instance, for both producer quality and consumer taste, they are likely to share a common factor with the overall recovered quality that is HS 6-digit specific.

Table 2: Product Quality and Consumer Taste Summary

Overall Quality: $\ln \phi_{jvt}^{V}$	Producer (Product) Quality: Exporter-specific $\ln \phi_{ivt}^{V}$	Consumer Taste: Importer-specific $\ln \phi_{jvt}^{V}$
Partial Correlation Coefficient	0.522	0.194
Variance Explained by Each Component individually	9.0%	
Variance Common to Each Component	69.2%	
Variance Explained in Total	78.2%	

Note: The sample periods include 2000, 2005, 2010 and 2015. Each variety is defined as exporter-HS 6-digit code for each importer in the given sector. Partial correlation coefficient measures the association between two variables, removing the confounding factor between the two.

In the variance decomposition exercise, producer quality and consumer taste jointly explain 78.2% of the total variance for variety quality, and other idiosyncratic shocks explain the rest 22.8%. Within the quality variance explained by consumer and producer, the statistics indicate that there exists an important common factor underlying these two margins that explain about 69% of the total variance, while consumer taste and producer quality, individually, only explain about 9% of the total. These statistics imply that producer quality is more responsible to variety quality on average globally.[24]

6.2 Global Trade Pattern Decomposition

We now apply our RCA decomposition method to investigate the importance of each mechanism underlying comparative advantage across countries and sectors. We start by decomposing the level and change of RCA in (15). Following Redding and Weinstein (2017), we adopt a variance decomposition introduced by Eaton, Kortum, and Kramarz (2004). Specifically, we study the contribution of each mechanism embedded in RCA by regressing each component on the overall value of RCA. Therefore, we will have three regressions with B-RCA as the independent variable, whose dependent variables are demand $\ln\left(RCA_{igt}\right)^{D}$, supply ($\ln\left(RCA_{igt}\right)^{S}$), and other factors unexplained by these two factors, respectively.

$$\begin{aligned}
\ln\left(RCA_{igt}\right)^{\mathbb{D}} &= \alpha_D + \beta_D \ln\left(RCA_{igt}\right) + \epsilon_{igt}^{D} \\
\ln\left(RCA_{igt}\right)^{\mathbb{S}} &= \alpha_S + \beta_S \ln\left(RCA_{igt}\right) + \epsilon_{igt}^{S} \\
\ln\left(RCA_{igt}\right)^{\mathbb{U}} &= \alpha_U + \beta_U \ln\left(RCA_{igt}\right) + \epsilon_{igt}^{U}
\end{aligned} \tag{26}$$

where each observations is exporter i, industry g and year t specific. By the property of OLS, we have $\beta_D + \beta_S + \beta_U = 1$, and the mangitude of each point estimates deliver the relative importance of each component of RCA. The overall supply ($\ln\left(RCA_{igt}\right)^{S}$) and demand factors ($\ln\left(RCA_{igt}\right)^{D}$) are defined as:

[24]Our finding is consistent with decomposition exercise by Di Comite, Thisse, and Vandenbussche (2014). Using Belgian firm-level data, they find the most of quality (55% of the variation in quantity) is explained by other factors (mostly producer quality), and consumer taste accounts for a smaller proportion (45%).

$$\ln\left(RCA_{igt}\right)^{\mathbb{D}} = \ln\left(RCA_{igt}^{\varphi^D}\right) + \ln\left[\frac{N_{igt}^{M}/N_{gt}^{EM}}{N_{it}^{MG}/N_{t}^{EMG}}\right] + \ln\left[RCA_{igt}\left(\mathbf{S},\tilde{\mathbf{w}},\mathbf{L}\right)\right]$$
$$\ln\left(RCA_{igt}\right)^{\mathbb{S}} = \ln\left(RCA_{igt}^{P}\right) + \ln\left(RCA_{igt}^{\varphi^S}\right) + \ln\left(RCA_{igt}^{S}\right) + \ln\left(RCA_{igt}^{N}\right)$$

Next, we investigate the relative importance of each margin within supply factors, and the OLS regressions are:

$$\begin{aligned}
\ln\left(RCA_{igt}^{P}\right) &= \alpha_P + \beta_P \ln\left(RCA_{igt}\right)^{\mathbb{S}} + \epsilon_{igt}^{P} \\
\ln\left(RCA_{igt}^{\varphi^S}\right) &= \alpha_{\varphi^S} + \beta_{\varphi^S} \ln\left(RCA_{igt}\right)^{\mathbb{S}} + \epsilon_{igt}^{\varphi^S} \\
\ln\left(RCA_{igt}^{S}\right) &= \alpha_S + \beta_S \ln\left(RCA_{igt}\right)^{\mathbb{S}} + \epsilon_{igt}^{S} \\
\ln\left(RCA_{igt}^{N}\right) &= \alpha_N + \beta_N \ln\left(RCA_{igt}\right)^{\mathbb{S}} + \epsilon_{igt}^{N}
\end{aligned} \tag{27}$$

We have $\beta_P + \beta_{\varphi^S} + \beta_S + \beta_N = 1$, and each coefficient tells us the relative importance of each mechanism within supply factors. Similarly, we regress the log change in each component on the overall log change in RCA to study its evolution pattern.

Table 3: Variance Decomposition Across Countries and HS 4-digit Sectors (2000-2015)

	% of Total Variance Explained	
	(1) Log Level	(2) Log Change
Overall $\ln RCA$		
$\ln\left(RCA_{igt}\right)^{\mathbb{S}}$	45%	41%
$\ln\left(RCA_{igt}\right)^{\mathbb{D}}$	23%	12%
$\ln\left(RCA_{igt}\right)^{\mathbb{U}}$	32%	47%
Supply Related $\ln\left(RCA_{igt}\right)^{\mathbb{S}}$		
$\ln(RCA_{igt}^{P})$	7%	5%
$\ln(RCA_{igt}^{\varphi^S})$	82%	86%
$\ln(RCA_{igt}^{N})$	5%	3%
$\ln(RCA_{igt}^{S})$	6%	6%

Note: Variance decomposition for the log level of RCA uses samples of 2000, 2005, 2010, and 2015. The log change in RCA use samples of 2005, 2010 and 2015, and the difference is relative to the year 2000.

In Table 3, we report the results of the decompositions for both RCA level in column (1), and changes in column (2). According to column (1), the supply and demand factors could explain more than two thirds of observed global RCA pattern across countries and years between 2000 and 2015. The remaining 32% RCAvariance captures other factors that are not explained by demand and supply, which can be due to other reasons such as the measurement errors in calculating real income, the measurement errors from approximating industry expenditure share, the idiosyncratic shocks to consumer perceived variety quality. We find that the supply factors of RCA are most important in accounting for the global trade pattern, explaining 45% of the total RCA variations, while demand factors are comparatively unimportant explaining about 23% of the global trade pattern.

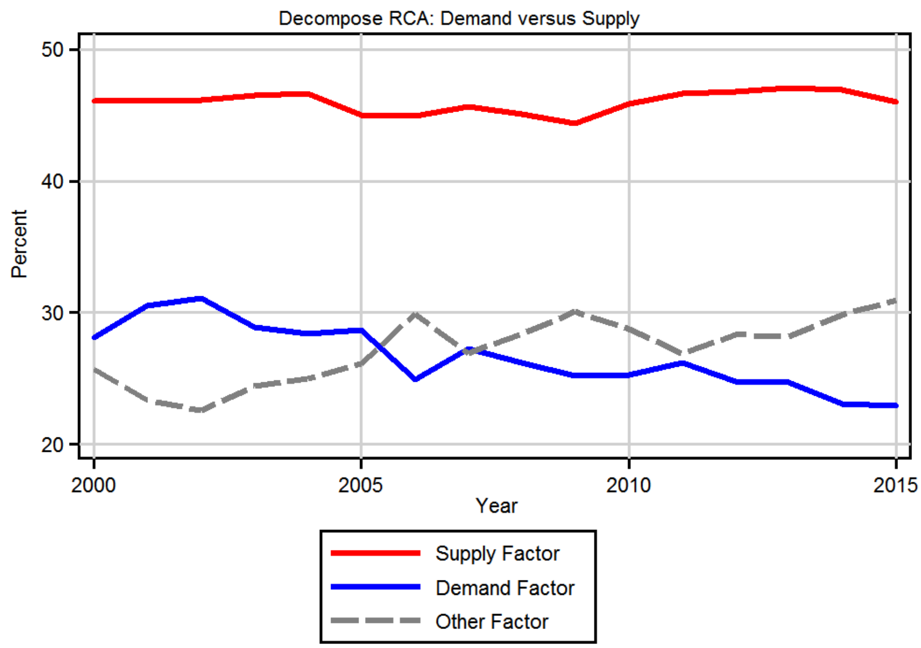


Figure 3: Decomposition of RCA: Demand and Supply Factor

In addition to the results from the pooled samples, we report the time-series results in Figure 3, which displays the annual percentage of RCA variance explained by demand (blue), supply (red), and other factors (grey). The proportions of global RCA variation explained by supply and demand factors remain stable, and we find supply factors remain persistently important across years. By contrast, there is a slight downward trend for the demand factors. The decomposition exercise also reveals that directly using RCA as the key supply factor in statistical inference could be misleading,[25] as the outcome variables, such as income measures like per capita GDP, could be also correlated with RCA via its demand margins. Such an issue will lead to the potential endogeneity of reverse causality when one tries to infer the causal impact of some variable based on RCA on some variables related to economic development.

[25] For instance, Bahar and Rapoport (2018) find that immigration from exporters of a given product promotes the host country's RCA of that product, and they attribute it as the productivity spillover from the immigrants. However, according to decomposition, it also implies changes in the host country's RCA could be due to the expansion via demand mechanisms, i.e., immigrants bring new markets to the host country.

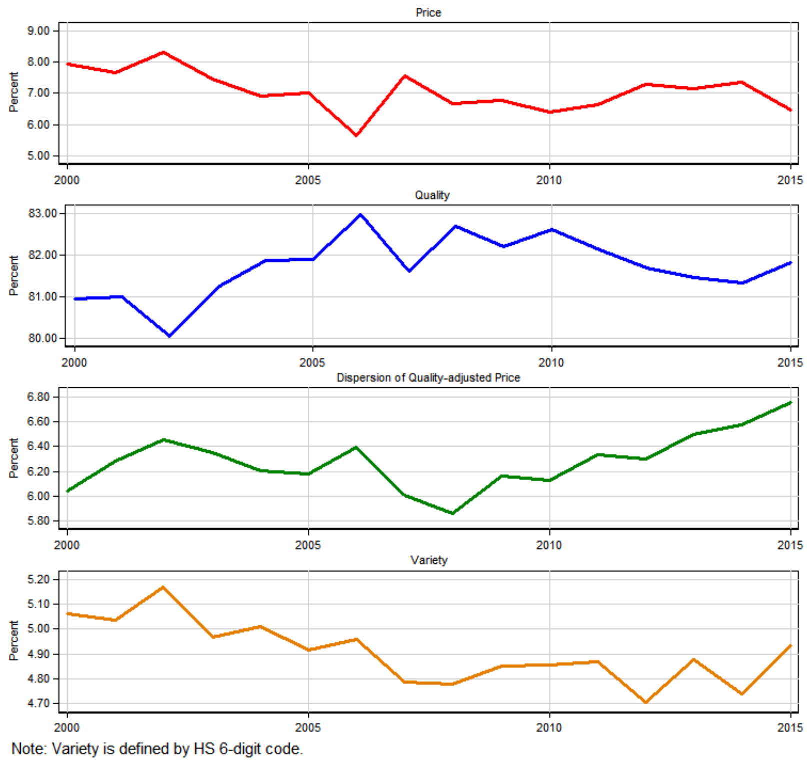


Figure 4: Decomposition of RCA: Within Supply Factors

Next, moving on to column (2) of Table 3, we find that average prices are comparatively unimportant in explaining patterns of trade associated with production. In the time series, average product prices account for 7% of the variation in $\ln\left(RCA_{igt}\right)^S$, i.e., 3.2% of the overall variation of $\ln\left(RCA_{igt}\right)$.[26] The unimportance for average variety prices in explaining trade pattern could be explained by the Heckscher-Ohlin model as indicated by Redding and Weinstein (2017). Trade equilibrium characterized by factor price equalization implies the relative variety prices are same across countries. By contrast, average product quality (producer quality) explains the most of production RCA, with a contribution of 37 percent for levels of RCA overall. As also pointed out by Redding and Weinstein (2017), the substantial difference also implies it is not obvious that the determinants of quality are the same as those of prices. For instance, a large literature in IO emphasizes the importance of sunk costs for product quality, such as Sutton (1991). Lastly, variety and the dispersion terms explain 2.3 percent and 2.7 percent of global trade pattern. Though being small in magnitude, they indicate the imperfect substitutability of varieties within sectors, as both the number of varieties and the dispersion terms also matter for the global pattern of trade.

In Figure 5, we report the pattern for global trade pattern evolution by focusing on the log changes of RCA relative to the year 2000. The summarizing statistics are provided in column (2) of Table 3.

[26]This finding is consistent with the firm-level study of Redding and Weinstein (2017) who also find product prices only explain 6.5% in their adjusted RCA.

Similar to the pattern for the level of RCA, we find supply factors in RCA are primarily responsible for the change of global trade pattern, explaining 41 percent of log change of RCA. By contrast, the demand RCA remains barely changed. Within the supplying factors, each mechanism is found to contribute to RCA evolution in the same way as they explain the variations for the level of RCA.

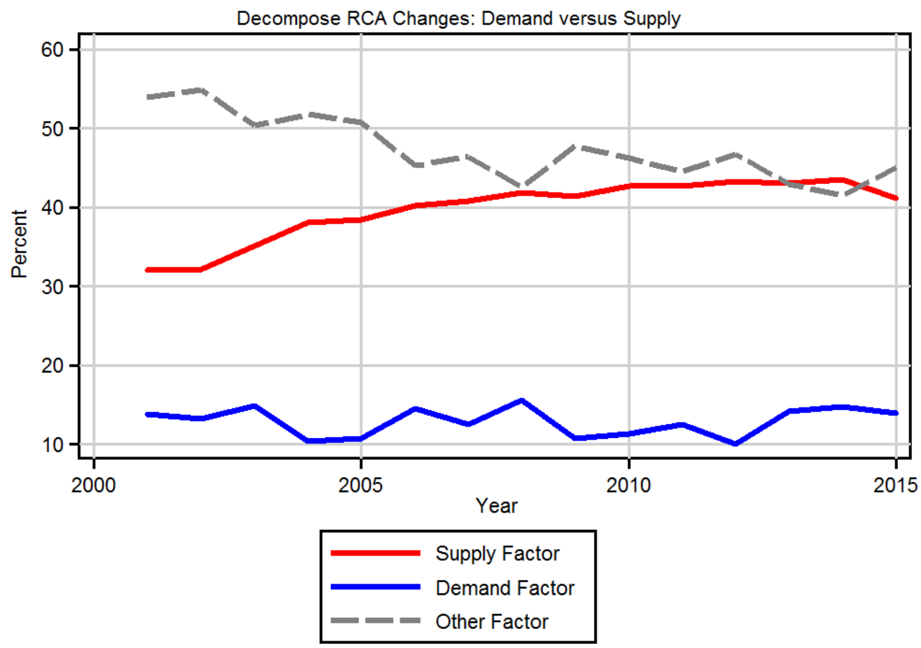


Figure 5: Decomposition of RCA Evolution: Demand and Supply Factor

Figure 6 displays the proportion of $\ln\left(RCA_{igt}\right)^S$ changes explained by each of the supply factor. Consistent with the findings in Redding and Weinstein (2017), we also find that the substantial churning in patterns of comparative advantage over time, as documented by recent empirical studies, are largely due to the changes in average product/producer quality, explaining about 86% of changes in $\ln\left(RCA_{igt}\right)^S$ (i.e., 35.5% of the changes in overall log RCA). By contrast, changes in export advantage associated with average prices remain stable across years, explaining about 5% of $\ln\left(RCA_{igt}\right)^S$ evolution. Notably, the proportion of variance explained by the variety and share dispersion are becoming greater in magnitude, indicating product differentiation becomes more critical for maintaining a healthy production advantage as also suggested by Hombert and Matray (2018).

To summarize, the results of this section highlight the role of demand and supply factors for explaining the world trade flows. Within supply factors, in line with Redding and Weinstein (2017), we also emphasize the role of imperfect substitutes in accounting for comparative advantage, as reflected by the dominant contribution from product quality to the production advantage. Therefore, a country's comparative advantage cannot be inferred solely from conventional measures of average prices. Instead, it also depends on the other non-conventional forces such as the number of varieties and product differentiation. In addition to Redding and Weinstein (2017), demand factors such as market access also affect a country's export performance and are also critical to explain the global trade pattern.

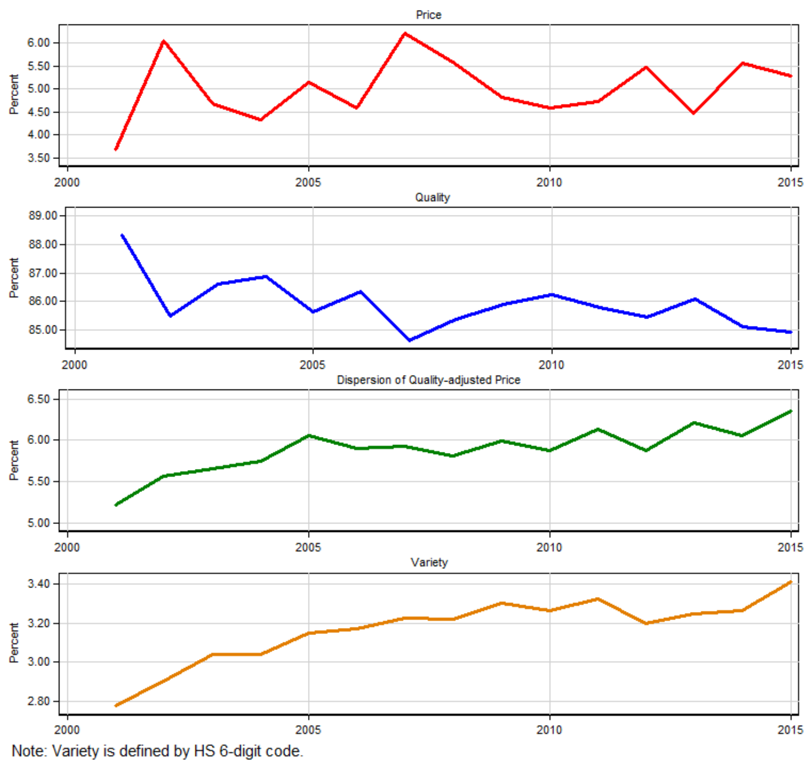


Figure 6: Decomposition of RCA Evolution: Within Supply Factors

6.3 Comparative Advantage and Economic Complexity

Recently, though, the introduction of measures of "economic complexity" as charted by Hidalgo and Hausmann (2009); Hausmann and Hidalgo (2011) has expanded our ability to quantify a country's productive structure. It uses bipartite international trade networks to characterize the economic complexity of a country based on the capabilities implied by the products it exports. The Economic Complexity Index (ECI) has received wide attention because it is highly predictive of future economic growth Hidalgo and Hausmann (2009); Hausmann and Hidalgo (2011), and national inequality Hartmann, Guevara, Jara-Figueroa, Aristarán, and Hidalgo (2017). As staggering as its explanation power is, little has been known about the underneath factors determining ECI. In this section, we apply our decomposition method to economic complexity to study its underlying channels, and to explore the contribution of each mechanism to various economic development.

Table 4: Economic Complexity Ranks (Top 15 Economies, 2015)

Country	(1) ECI Rank	(2) ECI^S Rank	(3) ECI^D Rank
Japan	1	2	1
Switzerland	2	1	3
Germany	3	4	29
South Korea	4	5	4
Finland	5	27	12
Sweden	6	16	10
the United States	7	10	26
Singapore	8	14	2
Austria	9	11	14
Norway	10	17	11
Czech Republic	11	18	6
Canada	12	47	9
Israel	13	40	19
United Kingdom	14	7	21
France	15	9	30

Note: We rank the countries that are listed in The Atlas of Economic Complexity (see https://atlas.media.mit.edu/en/rankings/country/eci/) between 2010 and 2015, and report the top 15 most complex economies based on their overall ECI.

Essentially, the ECI captures a country's productivity arising from the diversity of capabilities present in a country and their interactions (Hidalgo and Hausmann (2009)). Its construction rest on identifying a country's industries with comparative advantage according to Balassa RCA.[27] To study the mechanism of ECI, we are constructing ECI using each RCA component as appearing in (15), and each ECI will correspond to a particular specific mechanism. Table 4 reports the country ranks based upon different measures of complexity in the year 2015. According to column (1), Japan, Switzerland, and Germany are the top three economies with respect to the overall complexity level, and a majority of the listed economies are rich or OECD countries. Column (2) and column (3) display the corresponding country ranks based on ECI measures associated with supply and demand channels, respectively. We observe a heterogeneous pattern regarding the sources of economic complexity. For instance, Germany will be much lower ranked (29th) was the production structure solely determined by the foreign demand for German goods. It also implies that the sophistication of Germany's economy is mainly driven by factors on the production side such as good technology. By contrast, Singapore displays the opposite tendency, i.e., the economic complexity level of Singapore is more likely to be associated with market access, e.g., selling to big or rich countries.

Figure 7 displays the relationship between complexity measures associated with the demand and supply mechanisms. The coefficient of regressing ECI^Son ECI^D is significantly positive, implying countries that have better conditions on goods production are also like to have better market access. Meanwhile, there is notable heterogeneity in ECI^S among countries with different level of ECI^D, indicating the complexity measures based on demand and supply factors captures some different aspects of ECI.

[27] The method for constructing ECI is provided in Appendix D, and more details could refer to Hidalgo and Hausmann (2009).

Next, we compare the bivariate relationships between different measures of ECI constructed using each margin within supply facotr in Table 8.[28] According to panel (a), the quality ECI has a strong and significant negative correlation with price ECI, indicating countries capable of producing goods of better quality are likely to incur higher costs thus lose the export advantage from offering goods with cheaper prices.[29] In panel (b), countries producing high-quality goods are also more diversified in production within sectors, as reflected by the positive and significant correlation between $ECI^S_{quality}$and $ECI^S_{variety}$. Lastly, panel (c) shows that the negative relationship between the $ECI^S_{quality}$ and $ECI^S_{dispersion}$ is noticeably weaker and not significant statistically.

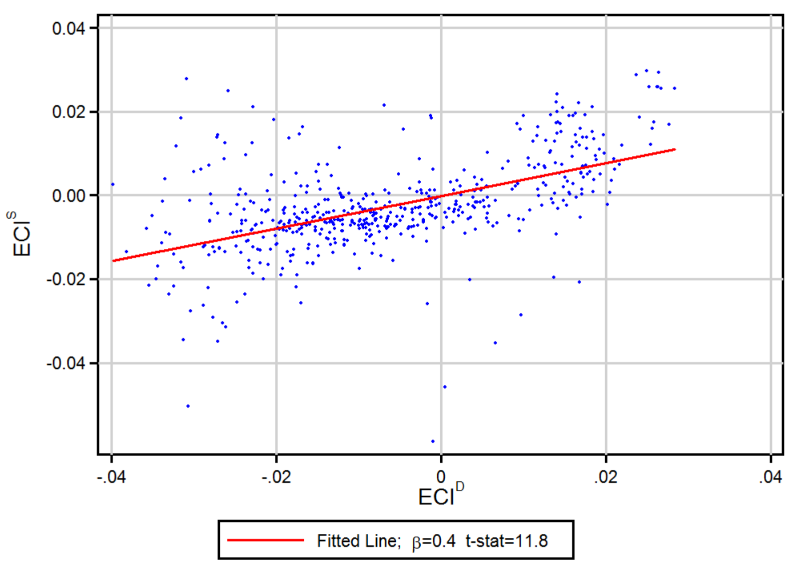


Figure 7: Relationship between ECI^S and ECI^D (2000-2015)

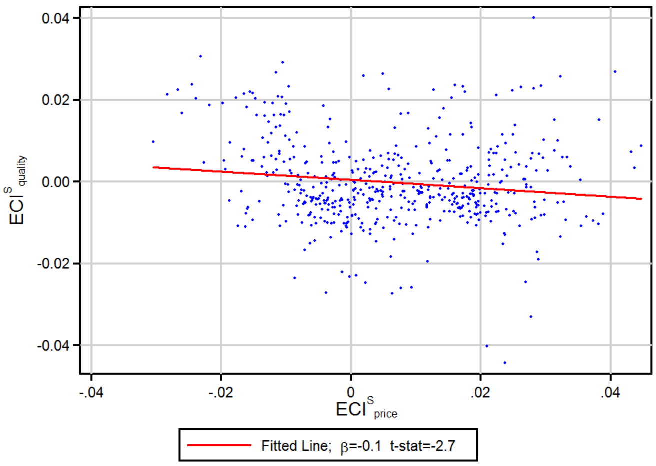
Given ECIs implied by each channel, with regression study, we explore which mechanism (s) determine a country complexity. Table 5 summarizes the contribution of each mechanism. The dependent variable is the original ECI constructed following Hidalgo and Hausmann (2009), and the independent variables are the ECI indexes created using various RCA components.[30] Standard errors are clustered at the country level. In the first three columns, we study the overall relevance of broad mechanisms, namely, supply and demand factors, and other unexplained RCA. Column (1) to (3) differ in regression specifications. The point estimates reveal that export advantage arising

[28]Table 8 uses the ECI measures for four periods, namely 2000, 2005, 2010, and 2015.

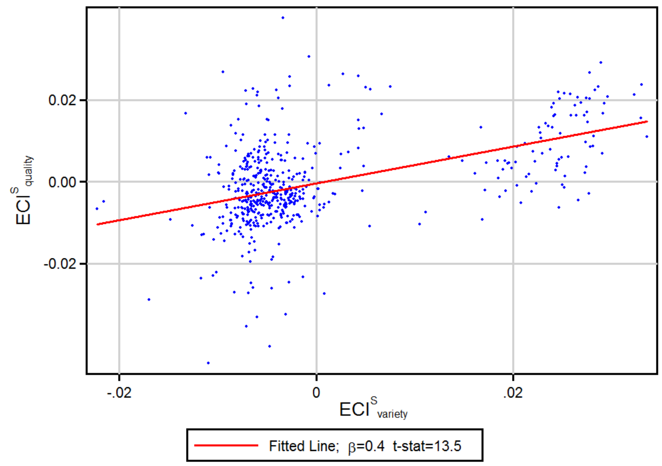
[29]This finding is consistent with the evidence discovered at the firm level (Fajgelbaum, Grossman, and Helpman (2011)) and country level (Kugler and Verhoogen (2011)).

[30]For instance, ECI^D refer to the ECI that is constructed using all demand factor $\ln\left(RCA_{igt}\right)^{\mathbb{D}}$; ECI^{Un} is the ECI constructed using the RCA component not related to demand and supply factors; among supply factors, ECI^S_{price} stands for the ECI constructed using $\ln\left(RCA^P_{igt}\right)$.

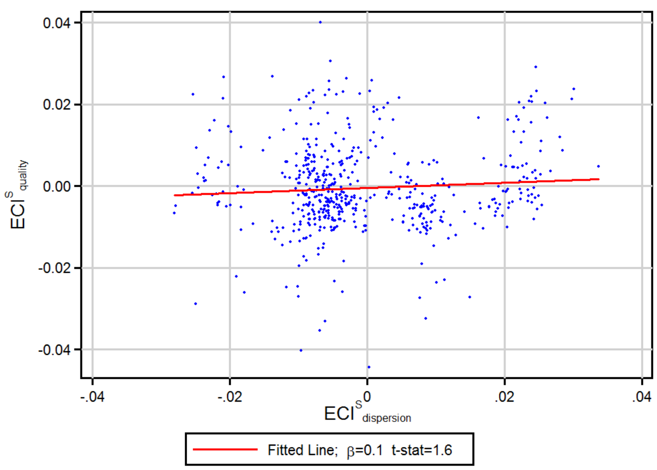
from both supply and demand channels as characterized in decomposition (15) are important in explaining the overall economic complexity of a country, as they are all significantly positive through different specifications. The ECI associated with the demand (ECI^D) has a greater impact on shaping the overall ECI compared to that associated with the supply factors (ECI^S). Notably, production advantage driven by some unknown reasons rather than supply and demand (ECI^{Un}) is not found to have significant explanatory power.



(a) $ECI^S_{quality}$ and ECI^S_{price}



(b) $ECI^S_{quality}$ and ECI^S_{price}



(c) $ECI^S_{quality}$ and $ECI^S_{dispersion}$

Figure 8: Economic Complexity Associations: Goods Supply Mechanisms (2000-2015)

In column (4) to (6), we further investigate the detailed channels among supply factors, where we repeat the regression but replacing ECI^S with the constructed ECI associated with average prices, producer quality, variety, and share dispersions. After controlling for year and country fixed effects, we find the more advantage countries have in charging competitive price, producing higher quality, and offering more variety, are more likely to be more complex in production space. Product differentiation as captured by the share dispersion ($ECI^S_{dispersion}$) is not found to associated with economic complexity overall. Among various supply factors, producer quality $ECI^S_{quality}$, and product diversification within an industry as captured by $ECI^S_{variety}$ are most important in the formation of economic complexity for a country, whose point estimates are about three times larger than that of average prices in the fixed effect models. [31]

Table 5: Economic Complexity Index by Margins

Dep var: ECI	(1) OLS	(2) FE	(3) FE	(4) OLS	(5) FE	(6) FE
ECI^D	0.327***	0.116***	0.079**	0.309***	0.125***	0.108***
	(0.030)	(0.035)	(0.036)	(0.035)	(0.030)	(0.033)
ECI^S	0.281***	0.058***	0.057***			
	(0.036)	(0.016)	(0.015)			
ECI^{Un}	-0.029	-0.009	-0.009	-0.031**	-0.009	-0.009*
	(0.018)	(0.007)	(0.007)	(0.015)	(0.005)	(0.005)
Breakdown of Supply Factors						
ECI^S_{price}				-0.010	0.036***	0.035***
				(0.027)	(0.012)	(0.012)
$ECI^S_{quality}$				0.294***	0.106***	0.117***
				(0.037)	(0.021)	(0.022)
$ECI^S_{variety}$				0.132**	0.146**	0.138**
				(0.055)	(0.057)	(0.056)
$ECI^S_{dispersion}$				0.001	0.004	0.002
				(0.019)	(0.007)	(0.007)
Observations	1,567	1,565	1,565	1,545	1,540	1,540
R-squared	0.532	0.933	0.935	0.623	0.942	0.943
Country FE	-	Y	Y	-	Y	Y
Year FE	-	-	Y	-	-	Y

Note: Standard errors are clustered at country level, and are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Taken together, the results of this section highlight the role of micro-mechanisms for a country's complexity. It is determined by the export advantage arising from both supply and demand channels. Empirically, we find demand factors play a slightly more important role in ECI. Within supply components, we find it is the producer quality and product diversification within sectors that are the most important to a country's overall ECI. Our findings open the black box of ambiguous ECI, and

[31] We also report the results of standardized ECI in Appendix G. In case of using standardized ECI, though export advantage unrelated to supply and demand becomes negatively correlated with overall ECI, both supply and demand factors remain the key factors to explain the overall ECI. Similarly, producer quality, product diversification within sectors, and average prices remain the top three most important factors for a country's overall economic complexity.

provide the rationale that reveals why ECI behaves similarly to product quality (Sutton and Trefler (2016)) in explaining the income growth.[32]

6.4 ECI and Economic Growth

In this section, we examine the contributions of each mechanism embedded in ECI to the growth of per capita GDP across countries. We adopt the simple version of the specification used in Hidalgo and Hausmann (2009):

$$\ln PGDP_i\left(t+\Delta t\right) = a + b_1 PGDP_i\left(t\right) + b_2 ECI_i\left(t\right) \tag{28}$$

where $PGDP_i(t)$ is per capita GDP for country i in year t, and $ECI_i(t)$ is the Economic Complexity Index of country i in year t (in practice, we substitute ECI with various ECI measures that are constructed using alternative RCA margins). We use the growth rates of per capita GDP in the different time span as the dependent variable. In the regression, we also control for the income level of the initial period.[33]

Table 6 reports the results using the 5-year growth rate of per capita GDP across countries in the world. In column (1), we replicate the same regression specification as in Hidalgo and Hausmann (2009) . Consistent with the previous findings, the point estimates of initial average income and ECI are significantly negative (-0.028) and positive (1.230), indicating the growth rate of average income is faster when a country has a lower level in initial average income and is more complex. Moving to column (2), we substitute the original ECI by ECI^D and ECI^S, which are constructed using $RCA_{igt}^{\mathbb{D}}$ and $RCA_{igt}^{\mathbb{S}}$. The alternative two ECI measures reflect the economic complexity associated with export advantages arising from the demand and supply mechanisms, respectively. [34] We find the point estimate of ECI^D is insignificant, while, in contrast, the estimate of ECI^S is significantly positive. The results imply that the majority of the impacts of economic complexity on income growth are from the supply channels, which have something to do with lowering marginal costs, improving product quality, expanding goods variety and raising product differentiation, rather than just trading with richer or bigger customers. This finding also addresses the criticism about potential bias in using complexity indexes to predict income growth, which arises from the concern that ECI computation may contain income information, and there may exist a spurious positive relationship between the two variables by construction. In column (3), we further split the overall supply factor into the contributions of a number of different micro-mechanisms, and investigate their impacts on growth. For the 5-year income growth rate, we find a rise in complexity channeled through competitive pricing and product diversification within sectors play the most crucial roles in boosting the economic

[32]It is also related to similar findings for technology sophistication (Lall, Weiss, and Zhang (2006)), as product quality is considered to require more technology input.

[33]The sample periods used in regression studies are 2000 to 2010 for 5-year growth, 2000-2008 for 7-year growth, and 2000-2005 for 10-year growth rate, respectively.

[34]As implied in Table 5 that the production advantage driven by some unknown reasons rather than supply and demand (ECI^{Un}) is not found to have significant explanatory power in overall ECI, we do not include it in the regression study.

growth in the short run, with the point estimates being 2.03 and 0.89, respectively. Improvement in producer quality or product differentiation is found less relevant to the income growth in this five-year span. Comparing the R-squared of second and third columns to our reference specification, we find that models incorporating different margins provide a better prediction of income growth than the reference specification using ECI alone.

Table 6: Economic Growth and Economic Complexity (5 Years)

	(1)	(2)	(3)
$PGDP(t)$	-0.028***	-0.029***	-0.022***
	(0.003)	(0.003)	(0.003)
ECI	1.230**		
	(0.526)		
ECI^D		0.763	1.505***
		(0.516)	(0.478)
ECI^S		1.379**	
		(0.616)	
Breakdown of Supply Factors			
ECI^S_{price}			2.034***
			(0.341)
$ECI^S_{variety}$			0.886*
			(0.460)
$ECI^S_{quality}$			-0.018
			(0.530)
$ECI^S_{dispersion}$			0.120
			(0.251)
Observations	1,397	1,212	1,212
R-squared	0.113	0.125	0.151

Note: The dependent variable is 5-year growth rate of per capita GDP, i.e., $lnPGDP(t + \Delta t)$. Sample periods in regression are 2000 to 2010. Robust standard errors are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

We now study the contribution of each margin to the economic growth in the longer run. The estimation target is the response of the alternative ECI measures were it applied in the growth rate of income in the longer periods. Table 7 and 8 report the results where we use 7-year and 10-year growth rate as the dependent variables. In the column (1) of each Table, the positive point estimates of ECI again confirm the contribution of economic complexity to the income growth for longer periods. Also consistent with Table 6, the point estimate of ECI^D remains insignificant in both tables, implying economic complexity interacts with economic growth mainly through the production structure rather than through the consumption structure. Among the supply factors, different from the previous pattern, the impact of product diversification within sectors now become unimportant to growth, as the coefficients of $ECI^S_{variety}$ become insignificant in Table 7 and 8. By contrast, product quality is found highly influential to per capita income growth, especially in the longer time periods. For example, a one standard deviation increase in $ECI^S_{quality}$ is associated with a rise in annual income growth rate by 0.2% and 0.3% in the time span of 7 and 10 years, respectively.[35] Cost advantage (ECI^S_{price}) remains an active channel to promote growth also in the long run. In addition, we find the magnitude of ECI^S_{price} coefficient is greater than that of $ECI^S_{quality}$ through Table 6 and 8 , implying productivity may influence growth more than quality in the long run.

Table 7: Economic Growth and Economic Complexity (7 Years)

	(1)	(2)	(3)
$PGDP(t)$	-0.043***	-0.043***	-0.034***
	(0.004)	(0.004)	(0.004)
ECI	2.098***		
	(0.759)		
ECI^D		0.661	1.443*
		(0.823)	(0.760)
ECI^S		2.313**	
		(0.987)	
Breakdown of Supply Factors			
ECI^S_{price}			2.794***
			(0.533)
$ECI^S_{variety}$			0.699
			(0.665)
$ECI^S_{quality}$			1.835**
			(0.857)
$ECI^S_{dispersion}$			0.184
			(0.343)
Observations	1,105	940	940
R-squared	0.154	0.168	0.200

Note: The dependent variable is 7-year growth rate of per capita GDP, i.e., $lnPGDP(t+\Delta t)$. Sample periods in regression are 2000 to 2008. Robust standard errors are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

To summarize, the results of this section highlight the channel of supply factors through which economic complexity boosts per capita income. We find the positive impact of complexity on growth is associated with lowering marginal costs, improving product quality, and expanding goods variety within sectors, while not relevant for product differentiation. Among these supply mechanisms, product diversification within sectors only matters in the short run, while product quality is found crucial in the long run. Comparing the estimates obtained from using the income growth rate of different time horizon as the outcome variables, we find greater production sophistication (higher

[35] Our empirical finding is also consistent with Sutton and Trefler (2016) who find advances in wealth are associated with changes in quality capabilities.

ECI) is more influential for economic growth in the long run. Complexity improvement associated with productivity change is more important than other mechanisms such as quality upgrading.[36]

Table 8: Economic Growth and Economic Complexity (10 Years)

	(1)	(2)	(3)
$PGDP(t)$	-0.061***	-0.062***	-0.046***
	(0.006)	(0.007)	(0.006)
ECI	3.175***		
	(1.088)		
ECI^D		0.528	1.801
		(1.348)	(1.278)
ECI^S		3.355**	
		(1.626)	
Breakdown of Supply Factors			
ECI^S_{price}			4.846***
			(0.801)
$ECI^S_{variety}$			0.091
			(1.158)
$ECI^S_{quality}$			3.416***
			(1.280)
$ECI^S_{dispersion}$			-0.301
			(0.512)
Observations	703	557	557
R-squared	0.192	0.224	0.292

Note: The dependent variable is 10-year growth rate of per capita GDP, i.e., $lnPGDP(t+\Delta t)$. Sample periods in regression are 2000 to 2005. Robust standard errors are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Causal Inference: A Structural IV Approach

Though ECI measures correlate strongly with income levels and predict subsequent economic growth with good statistical precision, previous identification strategy does not allow one to make causal inference with respect to the relationship between economic complexity and average income. As also demonstrated in (15), the overall RCA genetically bear the relationship with the average income across countries, which is captured by $\ln\left[RCA_{igt}\left(\mathbf{S}, \tilde{\mathbf{w}}, \mathbf{L}\right)\right]$. The ECI constructed based on B-RCA would be associated with income growth positively, if the demand mechanisms promote production diversification (ECI) of the country, and if these countries share a common factor in income growth. In this case, the OLS estimator would overestimate the impact of ECI on income growth.[37] To find

[36]Regression results of using standardize ECI are provided in Appendix G, and details refer to Table A.4 to Table A.6. The main findings remain robust.

[37]On the other hand, one will underestimate the impact of ECI on income growth, if the demand mechanisms dampen production diversification of the country (ECI).

a suitable instrument for economic complexity, we rely on the structure of the model displayed in (15). Instead of using the RCA in full expression, we isolate the variations thatare solely explained by supply factors, and use the corresponding ECI based on them as the instrument variable.[38]

In Table 9, we report estimation results with OLS and 2SLS methods. The first two columns use the 5-year income growth rate as the outcome. Consistent with the conceptual framework, the 2SLS point estimate of ECI is more significant and more prominent in magnitude than that of OLS, which further confirms the importance of the supply channels underlying ECI. Such pattern remains robust to all regressions as displayed in columns (3) to (6). The underestimation of ECI on income growth also indicates the market-oriented export is likely to dampen the country's economic complexity and slow down the subsequent economic growth. Such implication also sheds lights on the middle-income trap: countries specializing in commodities whose sales heavily depend on aggregate demand/market size will be more likely to suffer from middle-income trap.[39] Based on the 2SLS point estimates, a one standard deviation increase in overall ECI leads the annual income growth rate to rise by 0.65%, 0.74% and 0.72% in the time span of 5, 7 and 10 years, respectively. [40]

Table 9: Causal Inference: Economic Growth and Economic Complexity Index

	5-Year Growth		7-Year Growth		10-Year Growth	
Dep var: $lnPGDP(t + \Delta t)$	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS
$PGDP(t)$	-0.028***	-0.036***	-0.043***	-0.053***	-0.061***	-0.076***
	(0.003)	(0.004)	(0.004)	(0.006)	(0.006)	(0.010)
ECI	1.230**	4.244***	2.098***	5.844***	3.175***	7.366***
	(0.526)	(1.012)	(0.759)	(1.452)	(1.088)	(2.277)
Observations	1,397	1,216	1,105	941	703	557
R-squared	0.113	0.092	0.154	0.141	0.192	0.206

Note: The dependent variable is growth rate of per capita GDP. Robust standard errors are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

6.5 ECI and Inequality

Not only being predictive of future economic growth, economic complexity has also been found associated with a country's average level of income inequality (Hartmann, Guevara, Jara-Figueroa,

[38]The model implied instrument variable of ECI passes both Stock-Yogo and Kleibergen-Paap weak IV test at 10% significance level for all regressions differing in income growth rate. For instance, the first-stage F statistics is 158 in the regression using 10-year income growth rate as the dependent variable.

[39]Our implication on the middle-income trap is supported by the evidence found in Satoru (2011), Lin and Treichel (2012) and Gill and Kharas (2015), who document that countries relying less on primary commodities will be less likely to be trapped.

[40]The regression results using standardized ECI are provided in Table A.7 of Appendix G. The main findings of this section remain robust: the 2SLS estimates of ECI coefficients are significantly positive, and bigger in magnitude than that of OLS.

Aristarán, and Hidalgo (2017)).[41] For this reason, in this part, we explore the relative importance of each mechanisms underlying ECI in explaining the observed global inequality pattern. We strictly follow the empirical specification used in Hartmann, Guevara, Jara-Figueroa, Aristarán, and Hidalgo (2017), where the dependent variable is the average Gini index every five years, and all the controls are the same variables sourcing from the same database as used by those authors.[42]

Table 10: Inequality and Economic Complexity (2000-2005-2010)

Dep var: $Gini(ct)$	(1)	(2)	(3)
ECI	-0.992*		
	(0.551)		
ECI^D		-1.643***	-1.019**
		(0.430)	(0.438)
ECI^S		0.190	
		(0.406)	
Breakdown of Supply Factors			
ECI^S_{price}			-0.038
			(0.263)
$ECI^S_{variety}$			-0.864**
			(0.349)
$ECI^S_{quality}$			0.011
			(0.474)
$ECI^S_{dispersion}$			-0.011
			(0.196)
Observations	237	209	209
R-squared	0.571	0.609	0.646

Note: Following Hartmann et al (2017), we use the average Gini Index in a 5-year window as dependent variable, e.g., the average Gini index between 2000 and 2005 is used as the dependent variable for year 2000 (same for year 2005 and 2010). We also include the same controls as Hartmann et al (2017): $\ln(PGDP)$, $\ln(PGDP)^2$, years of schooling, $\ln(Population)$, rule of law, corruption control, government effectiveness, political stability, regulatory quality, and voice and accountability. Standard errors are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

We report the point estimates of ECI in Table 10. Column (1) uses the original complexity measure, and ECI is a negative and significant predictor of income inequality.[43] In column (2), we report regression results using ECI measures based on the overall demand and supply mechanisms. The point estimate of ECI^D is significantly negative while that of ECI^S remains insignificant, indicating

[41] According to these authors, production diversification will affect occupation distribution, which implies that countries differing in production structure will have mixed responses to global competitions with respect to income distribution (inequality). Some other channels that economic complexity affects inequality include knowledge spillover by human capitals and political powers.

[42] The detailed description on variables is provided in Appendix E.

[43] Consistently to Hartmann, Guevara, Jara-Figueroa, Aristarán, and Hidalgo (2017), we also find ECI explains the largest percentage of the variance in income inequality after the effects of all other variables have been taken into account.

the demand channels primarily drive the negative and significant bivariate relationship between a country's change in economic complexity and its Gini coefficient. Countries, whose production structure/diversification is driven by foreign demands, are likely to have a lower unequal distribution in income. Though in column (3), among the supply channels, we find diversification within sectors is also negatively correlated with inequality, as captured by -0.864 point estimate for $ECI^S_{variety}$, we do not find clear evidence supporting that an improvement in economic complexity due to better conditions among supplied factors are associated with inequality reduction, on average.[44] Our results illustrate that the ability of an economy to distribute income is strongly correlated with a country's economic complexity, which is mostly driven by demand factors underlying its production structure.

7 Conclusion

Using the bipartite international trade networks to characterize the economic complexity has generated new insights into the patterns of economic growth across countries, while the existing research on trade and on economic complexity remain methodologically disconnected. For the first time, leveraging the most commonly-used model in international trade, our paper provides a structural method for exploring the determinants of economic complexity and investigating the sources of economic development, which naturally bridges these two fields and provides a venue for future application.

Theoretically, we use our framework to derive a model-consistent formula that splits the traditional Balassa RCA into the contributions of a number of different micro-mechanisms. The decomposition exhibits that how good a country is in exporting in an industry depends on her average prices, product/producer quality, product varieties, and variety differentiation, in addition to demand factors such as market access. Empirically, we find supply factors are more important in accounting for the global trade pattern that explains twice as much variance as the demand channels.

Finally, our application to economic complexity reveals that comparative advantage arising from both supply and demand factors are relevant to the economy's sophistication. However, only the improvement on the production side could promote income growth effectively, while policies to encourage market access could have a limited impact. Our results also suggest that countries specializing in commodities whose sales heavily depend on aggregate demand will less likely to climb the complexity ladder and be more likely to suffer from middle-income trap.

[44] The estimation results of using standardized ECI are provided in Table A.8 of Appendix G. The pattern discussed in the main text remains robust.

References

BAHAR, D., R. HAUSMANN, AND C. A. HIDALGO (2014): "Neighbors and the evolution of the comparative advantage of nations: Evidence of international knowledge diffusion?," *Journal of International Economics*, 92(1), 111–123.

BAHAR, D., AND H. RAPOPORT (2018): "Migration, knowledge diffusion and the comparative advantage of nations," *The Economic Journal*, 128(612), F273–F305.

BALASSA, B. (1965): "Trade liberalization and "revealed" comparative advantage," *The Manchester School*, 33(2), 99–123.

BALLANCE, R. H., H. FORSTNER, AND T. MURRAY (1987): "Consistency tests of alternative measures of comparative advantage," *Weltwirtschaftliches Archiv*, pp. 157–161.

BERTHOU, A., AND C. EMLINGER (2011): "The trade unit values database," *Economie internationale*, (4), 97–117.

BLONIGEN, B. A. (2015): "Industrial policy and downstream export performance," *The Economic Journal*, 126(595), 1635–1659.

BOWEN, H. P. (1983): "On the theoretical interpretation of indices of trade intensity and revealed comparative advantage," *Weltwirtschaftliches Archiv*, 119(3), 464–472.

BRODA, C., AND D. E. WEINSTEIN (2006): "Globalization and the Gains from Variety," *The Quarterly Journal of Economics*, 121(2), 541–585.

BUSSIÈRE, M., G. CALLEGARI, F. GHIRONI, G. SESTIERI, AND N. YAMANO (2013): "Estimating trade elasticities: Demand composition and the trade collapse of 2008-2009," *American economic Journal: macroeconomics*, 5(3), 118–51.

CHEN, N., AND L. JUVENAL (2016): "Quality, trade, and exchange rate pass-through," *Journal of International Economics*, 100, 61–80.

COSTINOT, A., D. DONALDSON, AND I. KOMUNJER (2011): "What goods do countries trade? A quantitative exploration of Ricardo's ideas," *The Review of Economic Studies*, 79(2), 581–608.

DI COMITE, F., J.-F. THISSE, AND H. VANDENBUSSCHE (2014): "Verti-zontal differentiation in export markets," *Journal of International Economics*, 93(1), 50–66.

EATON, J., S. KORTUM, AND F. KRAMARZ (2004): "Dissecting trade: Firms, industries, and export destinations," *American Economic Review*, 94(2), 150–154.

ERKEL-ROUSSE, H., AND D. MIRZA (2002): "Import price elasticities: reconsidering the evidence," *Canadian Journal of Economics*, 35(2), 282–306.

Fajgelbaum, P., G. M. Grossman, and E. Helpman (2011): "Income Distribution, Product Quality, and International Trade," *Journal of Political Economy*, 119(4), 721–765.

Feenstra, R. C. (1994): "New product varieties and the measurement of international prices," *The American Economic Review*, pp. 157–177.

Feenstra, R. C., and J. Romalis (2014): "International prices and endogenous quality," *The Quarterly Journal of Economics*, 129(2), 477–527.

Feenstra, R. C., M. Xu, and A. Antoniades (2017): "What is the Price of Tea in China? Towards the Relative Cost of Living in Chinese and US Cities," Discussion paper, National Bureau of Economic Research.

French, S. (2017): "Revealed comparative advantage: What is it good for?," *Journal of International Economics*, 106, 83–103.

Gill, I. S., and H. Kharas (2015): *The middle-income trap turns ten*. The World Bank.

Hartmann, D., M. R. Guevara, C. Jara-Figueroa, M. Aristarán, and C. A. Hidalgo (2017): "Linking economic complexity, institutions, and income inequality," *World Development*, 93, 75–93.

Hausmann, R., and C. A. Hidalgo (2011): "The network structure of economic output," *Journal of Economic Growth*, 16(4), 309–342.

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

Hausmann, R., J. Hwang, and D. Rodrik (2007): "What you export matters," *Journal of economic growth*, 12(1), 1–25.

Hausmann, R., and B. Klinger (2006): "Structural transformation and patterns of comparative advantage in the product space," Discussion paper, Center for International Development, Harvard University.

Hidalgo, C. A., and R. Hausmann (2009): "The building blocks of economic complexity," *Proceedings of the National Academy of Sciences*, 106(26), 10570–10575.

Hoen, A. R., and J. Oosterhaven (2006): "On the measurement of comparative advantage," *The Annals of Regional Science*, 40(3), 677–691.

Hombert, J., and A. Matray (2018): "Can innovation help US manufacturing firms escape import competition from China?," *The Journal of Finance*, 73(5), 2003–2039.

Hottman, C. J., S. J. Redding, and D. E. Weinstein (2016): "Quantifying the sources of firm heterogeneity," *The Quarterly Journal of Economics*, 131(3), 1291–1364.

KUGLER, M., AND E. VERHOOGEN (2011): "Prices, plant size, and product quality," *The Review of Economic Studies*, 79(1), 307–339.

LALL, S., J. WEISS, AND J. ZHANG (2006): "The "sophistication" of exports: a new trade measure," *World development*, 34(2), 222–237.

LEVCHENKO, A. A., AND J. ZHANG (2016): "The evolution of comparative advantage: Measurement and welfare implications," *Journal of Monetary Economics*, 78, 96–111.

LIN, J. Y. (2012): *New structural economics: A framework for rethinking development and policy*. The World Bank.

LIN, J. Y., AND V. TREICHEL (2012): *Learning from China's rise to escape the middle-income trap: a new structural economics approach to Latin America*. The World Bank.

REDDING, S. J., AND D. E. WEINSTEIN (2016): "A unified approach to estimating demand and welfare," Discussion paper, National Bureau of Economic Research.

——— (2017): "Aggregating from micro to macro patterns of trade," Discussion paper, National Bureau of Economic Research.

RODRIK, D. (2006): "What's so special about China's exports?," *China & World Economy*, 14(5), 1–19.

SATORU, K. (2011): "The middle-income trap from the viewpoint of trade structures," Discussion paper, IDE Discussion Paper.

SUTTON, J. (1991): *Sunk costs and market structure: Price competition, advertising, and the evolution of concentration*. MIT press.

SUTTON, J., AND D. TREFLER (2016): "Capabilities, wealth, and trade," *Journal of Political Economy*, 124(3), 826–878.

TOKARICK, M. S. (2010): *A method for calculating export supply and import demand elasticities*, no. 10-180. International Monetary Fund.

VOLLRATH, T. L. (1991): "A theoretical evaluation of alternative trade intensity measures of revealed comparative advantage," *Weltwirtschaftliches Archiv*, 127(2), 265–280.

YEATS, A. J. (1985): "On the appropriate interpretation of the revealed comparative advantage index: implications of a methodology based on industry sector analysis," *Weltwirtschaftliches Archiv*, 121(1), 61–73.

YU, R., J. CAI, AND P. LEUNG (2009): "The normalized revealed comparative advantage index," *The Annals of Regional Science*, 43(1), 267–282.

Appendix

Theoretical Appendix

A. Rewrite RCA

The RCA formula used in Hidalgo and Hausmann (2009); Hausmann and Hidalgo (2011) could be written as:

$$
\begin{aligned}
RCA_{igt} &\equiv \frac{\left(\sum_{j\in\Omega_{igt}^{M}} X_{jigt}\right) / \left(\sum\sum_{i,j\in\Omega_{gt}^{EM}} X_{jigt}\right)}{\left(\sum\sum_{j,g\in\Omega_{it}^{MG}} X_{jigt}\right) / \left(\sum\sum\sum_{i,j,g\in\Omega_{t}^{EMG}} X_{jigt}\right)} \\
&= \frac{\left[\left(\sum_{j\in\Omega_{igt}^{M}} X_{jigt}\right) / N_{igt}^{M} \times N_{igt}^{M}\right] / \left[\left(\sum\sum_{i,j\in\Omega_{gt}^{EM}} X_{jigt}\right) / N_{gt}^{EM} \times N_{gt}^{EM}\right]}{\left[\left(\sum\sum_{j,g\in\Omega_{it}^{MG}} X_{jigt}\right) / N_{it}^{MG} \times N_{it}^{MG}\right] / \left[\left(\sum\sum\sum_{i,j,g\in\Omega_{t}^{EMG}} X_{jigt}\right) / N_{t}^{EMG} \times N_{t}^{EMG}\right]} \\
&= \frac{\left[\mathbb{M}_{igt}^{M}\left(X_{jigt}\right) \times N_{igt}^{M}\right] / \left[\mathbb{M}_{gt}^{EM}\left(X_{jigt}\right) \times N_{gt}^{EM}\right]}{\left[\mathbb{M}_{it}^{MG}\left(X_{jigt}\right) \times N_{it}^{MG}\right] / \left[\mathbb{M}_{t}^{EMG}\left(X_{jigt}\right) \times N_{t}^{EMG}\right]} \\
&= \frac{\mathbb{M}_{igt}^{M}\left(X_{jigt}\right) / \mathbb{M}_{gt}^{EM}\left(X_{jigt}\right)}{\mathbb{M}_{it}^{MG}\left(X_{jigt}\right) / \mathbb{M}_{t}^{EMG}\left(X_{jigt}\right)} \times \frac{N_{igt}^{M} / N_{gt}^{EM}}{N_{it}^{MG} / N_{t}^{EMG}}
\end{aligned}
$$

Note that X_{jigt} used are strictly positive so that we could apply geometric mean to the definition of RCA.

B. RCA Decomposition

After using geometric mean to approximate the arithmetic mean of trade flows, we obtain the new formula:

$$
RCA_{igt} = \Xi_{igt}\left(\mathbb{S}_{jigt}^{E}\right) \times \Xi_{igt}\left(\mathbb{X}_{jgt}^{E}\right) \times \frac{N_{igt}^{M} / N_{gt}^{EM}}{N_{it}^{MG} / N_{t}^{EMG}} \tag{29}
$$

where $\mathbb{X}_{jgt}^{E}$ is the total purchase of country j on the imported goods in sector g, and $\mathbb{S}_{jigt}^{E}$ indicates the share of the purchase that comes from country i.

We next express the total imports $\mathbb{X}_{jgt}^{E}$ as import share times the total expenditure (including the purchase made domestically) in sector g, which could be further broken down into margins displayed below:

$$
\begin{aligned}
\mathbb{X}^{E}_{jgt} &= \mu^{G}_{jgt} \times \mathbb{X}_{jgt} \\
&= \mu^{G}_{jgt} \times S_{jgt} \times Y_{jt} \\
&= \frac{\left(\mathbb{P}^{G}_{jgt}\right)^{1-\sigma^{V}_{g}}}{\left(P^{G}_{jgt}\right)^{1-\sigma^{V}_{g}}} \times \frac{\left(P^{G}_{jgt}\right)^{1-\sigma^{G}}}{\left(P_{jt}\right)^{1-\sigma^{G}}} \times w_{jt} L_{jt},
\end{aligned} \tag{30}
$$

where, in the first equality, μ^{G}_{jgt} is the import share for country j in sector g ; in the second equality, S_{jgt} is the total expenditure share of sector g in country j, and Y_{jt} is the total expenditure (GDP) in country j which equals $w_{jt}L_{jt}$ under trade balance. In the last equality, we express the share terms in the price ratios by the nature of CES, i.e., $\mu^{G}_{jgt} = \left(\mathbb{P}^{G}_{jgt}\right)^{1-\sigma^{V}_{g}} / \left(P^{G}_{jgt}\right)^{1-\sigma^{V}_{g}}$ and $S_{jgt} = \left(P^{G}_{jgt}\right)^{1-\sigma^{G}} / \left(P_{jt}\right)^{1-\sigma^{G}}$.

Combining (11), (29), and (30), along with the $\mathrm{S}^{E}_{jigt} = \left(\mathbb{P}^{E}_{jigt}\right)^{1-\sigma^{V}_{g}} / \left(\mathbb{P}^{G}_{jgt}\right)^{1-\sigma^{V}_{g}}$, we could rewrite RCA_{igt} as

$$
\begin{aligned}
RCA_{igt} &= \Xi_{igt}\left(\left[\mathbb{P}^{E}_{jigt}\right]^{1-\sigma^{V}_{g}}\right) \times \Xi_{igt}\left(\left[P^{G}_{jgt}/P_{jt}\right]^{\sigma^{V}_{g}-\sigma^{G}}\right) \times \Xi_{igt}\left(\left[w_{jt}/P_{jt}\right]^{1-\sigma^{V}_{g}}\right) \times \Xi_{igt}\left(w_{jt}^{\sigma^{V}_{g}} L_{jt}\right) \times \frac{N^{M}_{igt}/N^{EM}_{gt}}{N^{MG}_{it}/N^{EMG}_{t}} \\
&= \Xi_{igt}\left(\left[\mathbb{P}^{E}_{jigt}\right]^{1-\sigma^{V}_{g}}\right) \times \Xi_{igt}\left(\left[S_{jgt}\right]^{\frac{\sigma^{V}_{g}-\sigma^{G}}{1-\sigma^{G}}}\right) \times \Xi_{igt}\left(\left[w_{jt}/P_{jt}\right]^{1-\sigma^{V}_{g}}\right) \times \Xi_{igt}\left(w_{jt}^{\sigma^{V}_{g}} L_{jt}\right) \times \frac{N^{M}_{igt}/N^{EM}_{gt}}{N^{MG}_{it}/N^{EMG}_{t}}
\end{aligned}
$$

In the last step, to breakdown the price index $\mathbb{P}^{E}_{jigt}$, we first take logs on both sides of the equation:

$$
\begin{aligned}
\ln\left(RCA_{igt}\right) = \left(1-\sigma^{V}_{g}\right) & \left[\frac{1}{N^{M}_{igt}} \sum_{h\in\Omega^{M}_{igt}} \ln\left(\mathbb{P}^{E}_{higt}\right) - \frac{1}{N^{EM}_{gt}} \sum \sum_{h,l\in\Omega^{EM}_{gt}} \ln\left(\mathbb{P}^{E}_{hlgt}\right)\right] \\
& - \left[\frac{1}{N^{MG}_{it}} \sum \sum_{h,k\in\Omega^{MG}_{it}} \left(1-\sigma^{V}_{k}\right) \ln\left(\mathbb{P}^{E}_{hikt}\right) - \frac{1}{N^{EMG}_{t}} \sum\sum \sum_{h,l,k\in\Omega^{EMG}_{t}} \left(1-\sigma^{V}_{k}\right) \ln\left(\mathbb{P}^{E}_{hlkt}\right)\right] \\
& + \ln\left[\Xi_{igt}\left(\mathbf{S}, \tilde{\mathbf{w}}, \mathbf{L}\right)\right] + \ln\left[\frac{N^{M}_{igt}/N^{EM}_{gt}}{N^{MG}_{it}/N^{EMG}_{t}}\right]
\end{aligned}
$$

and subsitute $\ln\left(\mathbb{P}^{E}_{jigt}\right)$ with equation (10), which yields the exact log-linear decomposition of RCA as:

$$\ln RCA_{igt} \approx \underbrace{\ln\left(RCA^{P}_{igt}\right)}_{\text{Average prices}} + \underbrace{\ln\left(RCA^{\varphi}_{igt}\right)}_{\text{Average quality \& Taste}} + \underbrace{\ln\left(RCA^{S}_{igt}\right)}_{\text{Dispersion quality-adjusted prices}} + \underbrace{\ln\left(RCA^{N}_{igt}\right)}_{\text{Variety}}$$

$$+ \overbrace{\ln\left[\frac{N^{M}_{igt}/N^{EM}_{gt}}{N^{MG}_{it}/N^{EMG}_{t}}\right]}^{\text{Number of "Customers"}} + \overbrace{\ln\left[RCA_{igt}\left(\mathbf{S}, \tilde{\mathbf{w}}, \mathbf{L}\right)\right]}^{\text{Average Market Size}} \tag{31}$$

Provided that, in section (5), we could decompose the recovered product quality as:

$$\ln\left(\varphi^{V}_{jivt}\right) = \underbrace{\ln\left(\varphi^{V}_{jvt}\right)}_{\text{taste}\varphi^{D}} + \underbrace{\ln\left(\varphi^{V}_{ivt}\right)}_{\text{quality } \varphi^{S}} + \epsilon_{jivt}$$

The decomposition of (31) could further be written as:

$$\ln RCA_{igt} \approx \overbrace{\underbrace{\ln\left(RCA^{P}_{igt}\right)}_{\text{Average prices}} + \underbrace{\ln\left(RCA^{\varphi^{S}}_{igt}\right)}_{\text{Average quality}} + \underbrace{\ln\left(RCA^{S}_{igt}\right)}_{\text{Dispersion quality-adjusted prices}} + \underbrace{\ln\left(RCA^{N}_{igt}\right)}_{\text{Variety}}}^{Supply-side\ factors}$$

$$\underbrace{+ \overbrace{\ln\left(RCA^{\varphi^{D}}_{igt}\right)}^{\text{Average Taste}} + \overbrace{\ln\left[\frac{N^{M}_{igt}/N^{EM}_{gt}}{N^{MG}_{it}/N^{EMG}_{t}}\right]}^{\text{Number of "Customers"}} + \overbrace{\ln\left[RCA_{igt}\left(\mathbf{S}, \tilde{\mathbf{w}}, \mathbf{L}\right)\right]}^{\text{Average Market Size}}}_{Demand-side\ factors}$$

The expressions of these components are provided below:

$$\ln\left(RCA^{N}_{igt}\right) = \left\{ \begin{array}{l} \left[\frac{1}{N^{M}_{igt}} \sum_{h\in\Omega^{M}_{igt}} \ln N^{V}_{higt} - \frac{1}{N^{EM}_{gt}} \sum\sum_{h,l\in\Omega^{EM}_{gt}} \ln N^{V}_{hlgt}\right] \\ -\left[\frac{1}{N^{MG}_{it}} \sum\sum_{h,k\in\Omega^{MG}_{it}} \ln N^{V}_{hikt} - \frac{1}{N^{EMG}_{t}} \sum\sum\sum_{h,l,k\in\Omega^{EMG}_{t}} \ln N^{V}_{hlkt}\right] \end{array} \right\}$$

$$\ln\left(RCA^{P}_{igt}\right) = \left\{ \begin{array}{l} \left(1-\sigma^{V}_{g}\right)\left[\frac{1}{N^{M}_{igt}} \sum_{h\in\Omega^{M}_{igt}} \mathbb{E}^{V}_{higt}\left[\ln P^{V}_{vt}\right] - \frac{1}{N^{EM}_{gt}} \sum\sum_{h,l\in\Omega^{EM}_{gt}} \mathbb{E}^{V}_{hlgt}\left[\ln P^{V}_{vt}\right]\right] \\ -\left[\frac{1}{N^{MG}_{it}} \sum\sum_{h,k\in\Omega^{MG}_{it}} \left(1-\sigma^{V}_{k}\right) \mathbb{E}^{V}_{hikt}\left[\ln P^{V}_{vt}\right] - \frac{1}{N^{EMG}_{t}} \sum\sum\sum_{h,l,k\in\Omega^{EMG}_{t}} \left(1-\sigma^{V}_{k}\right) \mathbb{E}^{V}_{hlkt}\left[\ln P^{V}_{vt}\right]\right] \end{array} \right\}$$

$$\ln\left(RCA^{\varphi^{S}}_{igt}\right) = \left\{ \begin{array}{l} \left(\sigma^{V}_{g}-1\right)\left[\frac{1}{N^{M}_{igt}} \sum_{h\in\Omega^{M}_{igt}} \mathbb{E}^{V}_{higt}\left[\ln \varphi^{V}_{ivt}\right] - \frac{1}{N^{EM}_{gt}} \sum\sum_{h,l\in\Omega^{EM}_{gt}} \mathbb{E}^{V}_{hlgt}\left[\ln \varphi^{V}_{lvt}\right]\right] \\ -\left[\frac{1}{N^{MG}_{it}} \sum\sum_{h,k\in\Omega^{MG}_{it}} \left(\sigma^{V}_{k}-1\right) \mathbb{E}^{V}_{hikt}\left[\ln \varphi^{V}_{ivt}\right] - \frac{1}{N^{EMG}_{t}} \sum\sum\sum_{h,l,k\in\Omega^{EMG}_{t}} \left(\sigma^{V}_{k}-1\right) \mathbb{E}^{V}_{hlkt}\left[\ln \varphi^{V}_{lvt}\right]\right] \end{array} \right\}$$

$$\ln\left(RCA^{\varphi^{D}}_{igt}\right) = \left\{ \begin{array}{l} \left(\sigma^{V}_{g}-1\right)\left[\frac{1}{N^{M}_{igt}} \sum_{h\in\Omega^{M}_{igt}} \mathbb{E}^{V}_{higt}\left[\ln \varphi^{V}_{hvt}\right] - \frac{1}{N^{EM}_{gt}} \sum\sum_{h,l\in\Omega^{EM}_{gt}} \mathbb{E}^{V}_{hlgt}\left[\ln \varphi^{V}_{hvt}\right]\right] \\ -\left[\frac{1}{N^{MG}_{it}} \sum\sum_{h,k\in\Omega^{MG}_{it}} \left(\sigma^{V}_{k}-1\right) \mathbb{E}^{V}_{hikt}\left[\ln \varphi^{V}_{hvt}\right] - \frac{1}{N^{EMG}_{t}} \sum\sum\sum_{h,l,k\in\Omega^{EMG}_{t}} \left(\sigma^{V}_{k}-1\right) \mathbb{E}^{V}_{hlkt}\left[\ln \varphi^{V}_{hvt}\right]\right] \end{array} \right\}$$

$$\ln\left(RCA^{S}_{igt}\right) = -\left\{ \begin{array}{l} \left[\frac{1}{N^{M}_{igt}} \sum_{h\in\Omega^{M}_{igt}} \left(\mathbb{E}^{V}_{higt}\left[\ln s^{V}_{vt}\right] - \ln\frac{1}{N^{V}_{higt}}\right) - \frac{1}{N^{EM}_{gt}} \sum\sum_{h,l\in\Omega^{EM}_{gt}} \left(\mathbb{E}^{V}_{hlgt}\left[\ln s^{V}_{vt}\right] - \ln\frac{1}{N^{V}_{hlgt}}\right)\right] \\ -\left[\frac{1}{N^{MG}_{it}} \sum\sum_{h,k\in\Omega^{MG}_{it}} \left(\mathbb{E}^{V}_{hikt}\left[\ln s^{V}_{vt}\right] - \ln\frac{1}{N^{V}_{hikt}}\right) - \frac{1}{N^{EMG}_{t}} \sum\sum\sum_{h,l,k\in\Omega^{EMG}_{t}} \left(\mathbb{E}^{V}_{hlkt}\left[\ln s^{V}_{vt}\right] - \ln\frac{1}{N^{V}_{hlkt}}\right)\right] \end{array} \right\}$$

where Ω^M_{igt} is the set of foreign importer purchasing from exporter i within sector g at time t; $N^M_{igt} = \left|\Omega^M_{igt}\right|$ is the number of elements in set Ω^M_{igt}. Similar definition applies to other Ω and N. Demand factor $\ln\left[RCA_{igt}\left(\mathbf{S}, \tilde{\mathbf{w}}, \mathbf{L}\right)\right]$ is defined as

$$\ln RCA_{igt}\left(\mathbf{S}, \tilde{\mathbf{w}}, \mathbf{L}\right) \equiv \ln\left[\boldsymbol{\Xi}_{igt}\left(\left[S_{jgt}\right]^{\frac{\sigma_g^V - \sigma^G}{1-\sigma^G}}\right) \times \boldsymbol{\Xi}_{igt}\left(\left[w_{jt}/P_{jt}\right]^{1-\sigma_g^V}\right) \times \boldsymbol{\Xi}_{igt}\left(w_{jt}^{\sigma_g^V} L_{jt}\right)\right]$$

where variables $\mathbf{S}$, $\tilde{\mathbf{w}}$ and $\mathbf{L}$ denote the vectors for sector shares, real income, and population of each country.

Empirical Appendix

C. Methods of Imputing Expenditure Shares for HS 4-digit Industry

To apply the decomposition method, we require two share terms that are not available in data,[45] namely, country j's expenditure share for industry g (HS 4-digit) S_{jgt}, and import share in industry g (HS 4-digit) μ^{G}_{jgt}. Instead, we infer these statistics by combined using United Nations Comtrade Database (UN Comtrade), and the National Accounts Main Aggregate Database (NAMAD) that provides the production breakdown by ISIC aggregate sector for 200 countries between 2000 to 2015.

We first classify ISIC aggregate sectors into tradable and non-tradable sectors, which is summarized in Table A.1. In fact, all HS codes used in UN Comtrade could be concorded to the tradable ISIC aggregate sectors, and none of the HS codes could be linked to any of the non-tradable sectors.

Table A.1: List of ISIC Aggregate Sectors

Description	ISIC Sectors
Tradable Sectors ($\tilde{g} \in \Omega^T$)	
Agriculture, Hunting, Forestry, Fishing	A and B
Mining and Utilities	C and E
Manufacturing	D
Other Activities	J-P
Non-tradable Sectors ($\tilde{g} \in \Omega^G \setminus \Omega^T$)	
Construction	F
Wholesale, Retail Trade, Restaurants and Hotels	G and H
Transport, Storage and Communication	I

Note: The aggregate ISIC industry classification follows the National Accounts Main Aggregate Database.

We then calculate import share for all countries, and ISIC aggregate sectors that are denoted as $\tilde{g}$:

$$\mu^{G,ISIC}_{j\tilde{g}t} = \frac{IM^{ISIC}_{j\tilde{g}t}}{Y^{ISIC}_{j\tilde{g}t} - EX^{ISIC}_{j\tilde{g}t} + IM^{ISIC}_{j\tilde{g}t}} \tag{32}$$

where $Y^{ISIC}_{j\tilde{g}t}$, $EX^{ISIC}_{j\tilde{g}t}$ and $IM^{ISIC}_{j\tilde{g}t}$ denote the total output, export and import for ISIC aggregate sector $\tilde{g}$ in country j and year t. Note that, we directly source outpout from data, while indirectly calculating the same industry's export and import combined using UN Comtrade data. For instance, we calulate the sectoral gross import $IM^{ISIC}_{j\tilde{g}t}$ as:

$$IM^{ISIC}_{j\tilde{g}t} = \alpha_{\tilde{g}t} \times IM_{jt}$$

[45] As the Harmonized Commodity Coding System (HS code) is designed for international trade statistics, there is no corresponding production or consumption statistics that is summarized based on HS code.

where IM_{jt} is the reported country total imports in NAMAD, and the import share term $\alpha_{\tilde{g}t}$ is calculated combined using trade flows from UN Comtrade and the crosswalk from HS code to ISIC sectors. Similarly, one could compute $EX^{ISIC}_{j\tilde{g}t}$. Given import shares as provided in (32), we keep this ratio the same for all HS 4-digit industries that belong to ISIC aggregate sector $\tilde{g}$:

$$\mu^{G}_{jgt} = \mu^{G,ISIC}_{j\tilde{g}t}, \quad \text{if } g \text{ is concord to } \tilde{g}$$

Next, we compute the expenditure share of an HS 4-digit industry (g) as:

$$S_{jgt} = \frac{\mathbb{X}^{E}_{jgt}/\mu^{G}_{jgt}}{\sum_h \mathbb{X}^{E}_{jht}/\mu^{G}_{jht}} \times \mu^{T,ISIC}_{jt} \tag{33}$$

where $\mathbb{X}^{E}_{jgt}$ denotes the total imports in country j and industry g, i.e., $\mathbb{X}^{E}_{jgt} = \sum_{i\neq j} X_{jigt}$; expenditure share on tradable sector $\mu^{T,ISIC}_{jt}$ is computed as

$$\mu^{T,ISIC}_{jt} = \frac{\sum_{\tilde{h}\in\Omega^T}\left(Y^{ISIC}_{j\tilde{h}t} - EX^{ISIC}_{j\tilde{h}t} + IM^{ISIC}_{j\tilde{h}t}\right)}{\sum_{\tilde{h}\in\Omega^G}\left(Y^{ISIC}_{j\tilde{h}t} - EX^{ISIC}_{j\tilde{h}t} + IM^{ISIC}_{j\tilde{h}t}\right)}$$

where Ω^G contains both tradable and non-tradable sectors, and $\Omega^T \subset \Omega^G$ denotes the tradable sectors. For non-tradable sectors, the import and exports would be zero, i.e., $EX^{ISIC}_{j\tilde{h}t} = IM^{ISIC}_{j\tilde{h}t} = 0$ for $\tilde{h} \in \Omega^G \setminus \Omega^T$.

D. Computation of Economic Complexity Index (ECI)

The construction method follows Hidalgo and Hausmann (2009); Hausmann and Hidalgo (2011). Country i is regarded as a significant exporter of sector g if its RCA is greater than one.[46] Further, the elements M_{ig} can be expressed as:

$$M_{ig} = \begin{cases} 1 & RCA_{ig} > 1 \\ 0 & \text{otherwise} \end{cases}$$

By construction, $\boldsymbol{M}_{ig}$ is a matrix with rows indicating different countries and columns different industries. Thus, an element $(i,\ g)$ is 1 if country i is a significant exporter of product g. Therefore, we can measure a country's diversity and a product's ubiquity by summing over rows and columns respectively:

$$Diversity_i = k_{i,0} = \sum_g M_{ig}$$

[46]By adopting one as the 'cutoff value' we follow the convention in the literature.

$$Ubiquity_g = k_{g,0} = \sum_i M_{ig}$$

For countries, we calculate the average ubiquity of products that it exports, the average diversity of the countries that make those products and so forth. For products, we calculate the average diversity of the countries that make them and the average ubiquity of the other products that these countries make. The calculation can be expressed by the recursion defined as:

$$k_{i,N} = \frac{1}{k_{i,0}} \sum_g M_{ig} k_{g,N-1} \tag{34}$$

$$k_{g,N} = \frac{1}{k_{g,0}} \sum_g M_{ig} k_{i,N-1} \tag{35}$$

Substituting (34) into (35), we derive:

$$\begin{aligned} k_{i,N} &= \frac{1}{k_{i,0}} \sum_g M_{ig} \frac{1}{k_{g,0}} \sum_j M_{jg} k_{j,N-2} \\ &= \sum_j k_{j,N-2} \sum \frac{M_{ig} M_{jg}}{k_{i,0} k_{g,0}} \end{aligned} \tag{36}$$

We then define

$$\tilde{M}_{ij} = \sum_p \frac{M_{ig} M_{jg}}{k_{i,0} k_{g,0}} \tag{37}$$

Equation (36) can be rewritten as:

$$k_{i,N} = \sum_j \tilde{M}_{ij} k_{j,N-2} \tag{38}$$

Equation (38) is satisfied when $k_{i,N}$ and $k_{i,N-2}$ are both one, which corresponds to the eigenvector of $\tilde{M}_{ij}$ associated with the largest eigenvalue. Since this eigenvector is a vector of ones, which doesn't provide information, we instead use the eigenvector associated with the second largest eigenvalue, which provides the largest variation in the system to measure complexity. Thus, a country's ECI based on normalizing the eigenvector is expressed as:

$$ECI = \frac{\vec{K} - Mean(\vec{K})}{StdDev(\vec{K})} \tag{39}$$

where $\vec{K}$ is the eigenvector of $\tilde{M}_{ij}$ associated with the second largest eigenvalue, $Mean(.)$ denotes the average and $StdDev(.)$ represents the standard deviation of this eigenvalue.[47] In practice, we use

[47] Analogously, the Product Complexity Index (PCI) is derived as $PCI = (\vec{Q} - Mean(\vec{Q}))/S.D(\vec{Q})$, where $\vec{Q}$ is the eigenvector of $\tilde{M}_{gh}$ associated with the second largest eigenvalue.

each component of RCA as displayed in (15) to compute the economic complexity index correspondingly.

E. Detailed Variable Description: Inequality and ECI

The specification studying the relationship between ECI and inequality in Hartmann, Guevara, Jara-Figueroa, Aristarán, and Hidalgo (2017) is provided below:

$$Gini_{ct} = \beta_0 + \beta_1 ECI + \beta_2 \ln PGDP_{ct} + \beta_3 (\ln PGDP_{ct})^2 + \beta_4 Schooling + \beta_5 \ln Pop_{ct} + Z'\gamma + \epsilon_{ct}$$

The pooled OLS empirical model regresses income inequality against economic complexity, a country's average income and its square, population, human capital, and other institutional controls such as rule of law, corruption control, government effectiveness, political stability, regulatory quality, and voice and accountability. The source of each variable refer to Table A.2. The coefficient of interest is β_1 in the original study. In practice, we will substitute ECI with various ECI measures that are constructed using alternative mechanisms, and study their contributions to income inequality. Because of the sparseness of several variables, especially the Gini data, we follow the original paper by using the average values of Gini coefficien between periods 2000-2005, 2005-2010, and 2010-2015, to measure the Gini index for the year 2000, 2005 and 2010, respectively.

Table A.2: Variable Description: Inequality and ECI

Variable Names	Data Source
Gini Coefficient ($Gini_{ct}$)	GINI EHII and GINI ALL Datasets
Economic Complexity (ECI_{ct})	By Construction
GDP per capita ($\ln PGDP_{ct}$)	World Bank: World Development Indicators
Population ($\ln Pop_{ct}$)	Same Above
Average Years of Schooling ($Schooling$)	Same Above
Institutional Controls (Z)	
Rule of Law	World Bank: Worldwide Governance Indicators
Corruption Control	Same Above
Political Stability	Same Above
Government Effectiveness	Same Above
Regulatory Quality	Same Above
Voice and Accountability	Same Above

F. Appendix Figures

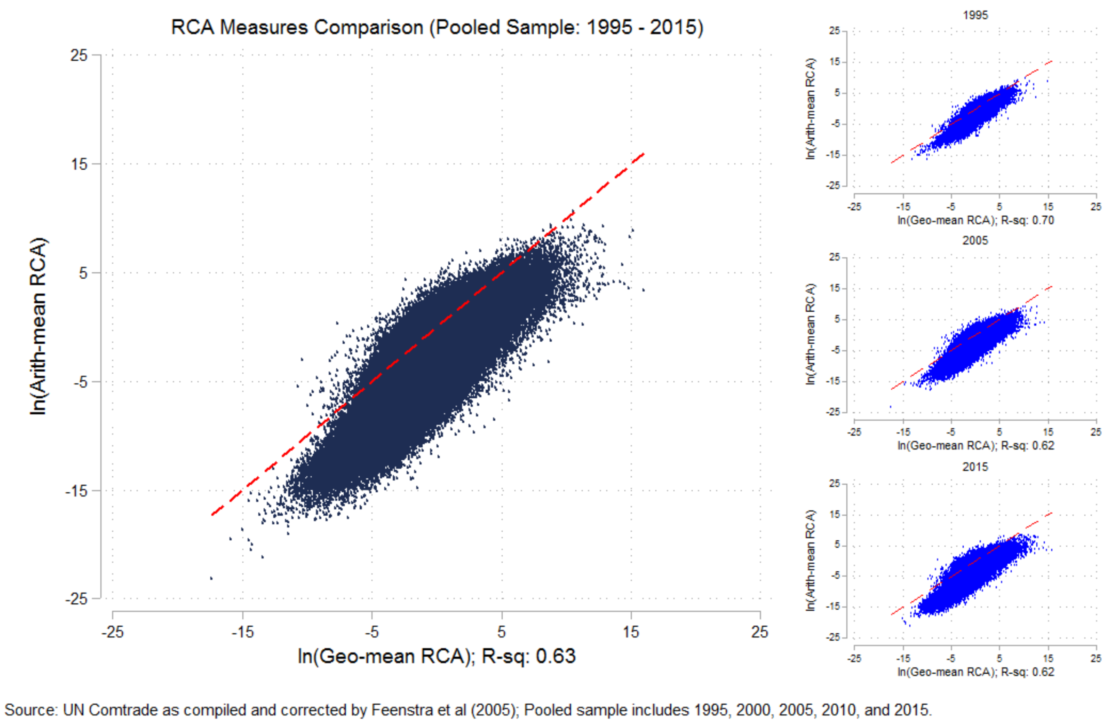


Figure A.1: RCA Measures Comparison (4-digit SITC Rev 2)

G. Appendix Tables

Table A.3: Economic Complexity Index by Margins

Dep var: ECI	(1) OLS	(2) FE	(3) FE	(4) OLS	(5) FE	(6) FE
ECI^{D}	0.475***	0.144***	0.097*	0.425***	0.174***	0.149***
	(0.046)	(0.055)	(0.057)	(0.054)	(0.051)	(0.055)
ECI^{S}	0.307***	0.054***	0.047***			
	(0.040)	(0.017)	(0.017)			
ECI^{Un}	-0.141***	-0.064**	-0.061**	-0.137***	-0.050**	-0.049**
	(0.037)	(0.029)	(0.027)	(0.035)	(0.021)	(0.020)
Breakdown of Supply Factors						
ECI^{S}_{price}				-0.036	0.055***	0.048***
				(0.042)	(0.018)	(0.017)
$ECI^{S}_{quality}$				0.319***	0.113***	0.108***
				(0.042)	(0.023)	(0.023)
$ECI^{S}_{variety}$				0.148**	0.157***	0.143**
				(0.061)	(0.060)	(0.061)
$ECI^{S}_{dispersion}$				0.036	0.054***	0.035***
				(0.029)	(0.012)	(0.010)
Observations	1,567	1,565	1,565	1,545	1,540	1,540
R-squared	0.534	0.927	0.936	0.627	0.936	0.943
Country FE	-	Y	Y	-	Y	Y
Year FE	-	-	Y	-	-	Y

Note: Standard errors are clustered at country level, and are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table A.4: Economic Growth and Standardized Economic Complexity (5 Years)

	(1)	(2)	(3)
$PGDP(t)$	-0.028***	-0.028***	-0.021***
	(0.003)	(0.003)	(0.003)
ECI	0.014***		
	(0.005)		
ECI^D		0.014*	0.021***
		(0.008)	(0.007)
ECI^S		0.009	
		(0.007)	
Breakdown of Supply Factors			
ECI^S_{price}			0.031***
			(0.005)
$ECI^S_{variety}$			0.013**
			(0.005)
$ECI^S_{quality}$			-0.002
			(0.006)
$ECI^S_{dispersion}$			-0.000
			(0.003)
Observations	1,397	1,215	1,215
R-squared	0.114	0.120	0.153

Note: The dependent variable is 5-year growth rate of per capita GDP, i.e., $lnPGDP(t+\Delta t)$. Sample periods in regression are 2000 to 2010. Robust standard errors are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table A.5: Economic Growth and Standardized Economic Complexity (7 Years)

	(1)	(2)	(3)
$PGDP(t)$	-0.044***	-0.043***	-0.034***
	(0.004)	(0.004)	(0.004)
ECI	0.028***		
	(0.008)		
ECI^D		0.014	0.024**
		(0.012)	(0.011)
ECI^S		0.019*	
		(0.010)	
Breakdown of Supply Factors			
ECI^S_{price}			0.043***
			(0.008)
$ECI^S_{variety}$			0.011
			(0.008)
$ECI^S_{quality}$			0.014
			(0.009)
$ECI^S_{dispersion}$			0.003
			(0.004)
Observations	1,105	940	940
R-squared	0.159	0.161	0.202

Note: The dependent variable is 7-year growth rate of per capita GDP, i.e., $lnPGDP(t + \Delta t)$. Sample periods in regression are 2000 to 2007. Robust standard errors are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table A.6: Economic Growth and Standardized Economic Complexity (10 Years)

	(1)	(2)	(3)
$PGDP(t)$	-0.061***	-0.062***	-0.046***
	(0.006)	(0.007)	(0.006)
ECI	0.034***		
	(0.012)		
ECI^D		0.008	0.028
		(0.020)	(0.020)
ECI^S		0.035**	
		(0.018)	
Breakdown of Supply Factors			
ECI^S_{price}			0.073***
			(0.012)
$ECI^S_{variety}$			0.003
			(0.014)
$ECI^S_{quality}$			0.034**
			(0.014)
$ECI^S_{dispersion}$			0.000
			(0.007)
Observations	703	557	557
R-squared	0.192	0.223	0.290

Note: The dependent variable is 10-year growth rate of per capita GDP, i.e., $lnPGDP(t + \Delta t)$. Sample periods in regression are 2000 to 2005. Robust standard errors are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table A.7: Causal Inference: Economic Growth and Standardized Economic Complexity Index

	5-Year Growth		7-Year Growth		10-Year Growth	
Dep var: $lnPGDP(t + \Delta t)$	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS
$PGDP(t)$	-0.028***	-0.033***	-0.044***	-0.051***	-0.061***	-0.074***
	(0.003)	(0.004)	(0.004)	(0.006)	(0.006)	(0.010)
ECI	0.014***	0.033***	0.028***	0.053***	0.034***	0.074***
	(0.005)	(0.010)	(0.008)	(0.014)	(0.012)	(0.024)
Observations	1,397	1,216	1,105	941	703	557
R-squared	0.114	0.107	0.159	0.157	0.192	0.209

Note: The dependent variable is growth rate of per capita GDP. Robust standard errors are reported in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table A.8: Inequality and Standardized Economic Complexity (2000-2005-2010)

Dep var: $Gini(ct)$	(1)	(2)	(3)
ECI	-0.009*		
	(0.005)		
ECI^D		-0.026***	-0.017**
		(0.007)	(0.007)
ECI^S		0.002	
		(0.005)	
Breakdown of Supply Factors			
ECI^S_{price}			-0.001
			(0.004)
$ECI^S_{variety}$			-0.008**
			(0.004)
$ECI^S_{quality}$			-0.003
			(0.005)
$ECI^S_{dispersion}$			-0.001
			(0.002)
Observations	237	209	209
R-squared	0.571	0.610	0.648

Note: Following Hartmann et al (2017), we use the average Gini Index in a 5-year window as dependent variable, e.g., the average Gini index between 2000 and 2005 is used as the dependent variable for year 2000 (same for year 2005 and 2010). We also include the same controls as Hartmann et al (2017): $\ln(PGDP)$, $\ln(PGDP)^2$, years of schooling, $\ln(Population)$, rule of law, corruption control, government effectiveness, political stability, regulatory quality, and voice and accountability. Standard errors are reported in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$.