Production Ability and Economic Growth*

Sebastián Bustos†
Harvard University
CID at Harvard University§

Muhammed A. Yıldırım ‡
Koç University
CID at Harvard University§

March 1, 2019

Abstract

Production is shaped by capability requirements of products and availability of these capabilities across locations. We propose a capabilities based production model and an empirical strategy to measure product sophistication and location's production ability. We apply our framework to international trade data, and employment data in the US, recovering measures of production ability for countries and cities, and sophistication of products and industries. We show that both country and city level measures have a strong correlation with income, and economic growth at different time horizons. Product sophistication is positively correlated with measures like education and training needed in the industry. Our model-based estimations also predict the diversification patterns through the extensive margin.

JEL Codes: O41, O47, O50, F10, F11, F14

*We are grateful to Ricardo Hausmann for insightful conversations, continuous encouragement and support. We would like to thank Carla Tokman and Daniel P. Stock for their comments. All errors are our own.

†Bustos: Harvard Kennedy School, 79 JFK Street, Cambridge, MA, 02138. E-mail: sebastian_bustos@hks.harvard.edu

‡Yıldırım: Koç Üniversitesi, Rumelifeneri Yolu, 34450 Sarıyer/İstanbul Turkey. E-mail: mayildirim@ku.edu.tr

§Center for International Development at Harvard University

1 Introduction

Economic development and growth is tightly linked with the industrial structure of a location (Marshall, 1890; Jacobs, 1969; Porter, 2003). Industries have many requirements that must be fulfilled locally in order to function competitively[1]. We call these required features *capabilities*, with those products requiring many different capabilities are deemed more *sophisticated*. Locations are successful in hosting these industries insofar as they can provide such capabilities or necessities. Locations that have a great variety of capabilities will therefore be more *able* for production. In this paper we propose a simple stochastic model, focusing on the capability requirements of products and offerings by locations, to measure the sophistication of products and abilities of locations. We then perform two empirical applications measuring production ability; across countries in the world, and cities in the United States. Our results show that countries and cities with higher production ability are expected to grow faster over time.

What are these capabilities? Let's turn to an example. Suppose a businessperson would like to open a new production facility. She needs to understand how local institutions work (whether and how property rights and rule of law are enforced in that location), and needs to obtain the formal permits to operate. She needs access to finance, and a contractor that can build up the physical space which, in turn, requires the construction materials. To run the plant, she needs access to a power grid, water distribution, energy lines and a sewerage system. She should be able to provide all the necessary inputs for production locally, or she needs to be able to buy and transport them from other locations, which requires a well-functioning customs mechanism and, again, a suitable infrastructure. Once the plant is ready to make the output, she still has to take her products to her customers. Customers may also want to visit the plant to inspect the standards. Therefore, well-functioning customs, ports and airports are crucial. Maybe the most important, the *productive knowledge* to make the operation possible needs to be brought together: she would need to assemble the necessary workforce, including managers, engineers, accountants, human resource specialists, and unskilled workers, to name a few. These workers need a place to stay, restaurants to eat, access to healthcare, and schools and daycare centers for their children. This list is by no means comprehensive; we have not yet touched upon any industry-specific requirements, for example. Locations, on the other hand,

[1] We chose the word location in order to refer to cities, metropolitan areas or countries.

must offer capabilities required by an industry to sustainably host it. Some of the capabilities are (non-rival) public inputs usually provided by governments, such as rule of law, roads or specific regulation. Other capabilities[2] are privately provided and organized by markets, raising the issue of positive and negative externalities. By increasing their capabilities, locations compete to become home to new industries. Therefore, the process of economic diversification can be thought as process of accumulating capabilities, a concept more complex that accumulating capital and labor. Importantly, many of these capabilities are used by multiple industries, making it possible for the burden of diversification to be shared between them.Consequently, the distribution of industries and products in the world is not random, and depends on the capability profile of products and locations.[3] However, as can be seen in the example above, listing all the capability requirements of industries or capability offerings of locations exhaustively might be unfeasible and inherently prone to mistakes. Properly measuring capabilities could be either an endless or impossible task. Nevertheless, the manifestation of capabilities are revealed in the products and where they are produced; products are embodiments of these requirement. Thus, products made within a location are indicators of capabilities present in that location.

In this paper, by imposing a nested structure on the capability requirements of products, we propose a method to estimate the extent of capabilities required by products, which we call *sophistication*, and a measure of a location's capability offerings, which we call *production ability*.[4] Particularly, we build on the model introduced by Hausmann and Hidalgo (2011), where the capability requirements of products and capability profiles of countries were matched using a Leontief-inspired production function. Based on the nestedness assumption, the relevant statistics to keep track of the product sophistication is the norm of its product capability-requirement vector. On the other hand, we can use the products as yardsticks, matching the level of a country's ability to the sophistication of the product that it makes with the highest norm. In the Leontief case, this corresponds to a threshold level for a country; any product whose sophistication is below this threshold could be produced by the country, any product with sophistication higher than the threshold will not be made by the

[2]We will assume that all capabilities are non-rivalrous

[3]There are well known economic theories that emphasize that the profile of products and locations shape the output process. For instance, in international trade, Hecksher-Ohlin proposes that factor endowment and input requirements explain the pattern of trade.

[4]In this context nested – in a strict interpretation – means that if we compare the capability requirements of any two products, one will be the subset of the other. In other words, the capabilities of less sophisticated product is nested under the capabilities required for the more sophisticated product for any product pair.

country.

In practice, however, the Leontief production function and perfect nestedness assumption are too restrictive for conducting an empirical exploration. To allow for substitution between capabilities, and also for differential capability accumulation patterns[5], we adapt a probabilistic approximation to the Leontief production. Moreover, not all products will be equally sensitive to the production ability of locations. For instance, a natural resource based product –like copper– will be made in relatively few places, but the countries that produce copper do so because of geographical luck rather than capability accumulation. Hence, copper is expected to have less sensitivity towards the productive ability differences across countries. Yet, copper is still a relatively specific product because it is only produced in few places. In contrast, more knowledge intensive products will exhibit both high levels of specificity and high levels of sensitivity. Therefore, instead of imposing a hard threshold associated with sophistication –as in the Leontief production function– we relax this assumption and introduce two different facets of product sophistication: *product specificity* and *product sensitivity*. To be explicit, first, there will be a *specificity* level associated with a product, and as production ability comes closer to and passes this level, the likelihood of making this product will increase. Second, we assume that products differ in their *sensitivity* to capabilities, meaning that the some products are more sensitive to the ability differences between locations than others. In practice, the likelihood of making a product depends on all three features; production ability of the location, the product specificity and the product sensitivity. A stochastic approximation to a Leontief function with these features can be captured by a logit functional form. With this specification, we can estimate the product sophistication and country ability measures by maximizing the overall likelihood function.

An analogy may help clarify the concepts behind our framework. Assume that production happening in a location is analogous to students taking a test (e.g. mathematics exam). Whether a location succeeds in making a product is equivalent to answering a question correctly. Students with higher ability are expected to answer more questions correctly. But not all question are alike. Some questions require higher math ability, and those students with more knowledge and preparation are more likely to answer them correctly. Hence, these questions are more *sensitive* to the math ability of students. Among the questions sensitive to math ability, some will be very difficult,

[5]other than linear path implied by the nestedness

demand high-ability and look for very *specific* knowledge. But specificity sometimes could be unrelated to demand for higher ability. For instance, in the math test we can think of a question asking the tenth digit of π. Possibly some students memorized the digits of π will get this question right, but this might not be necessarily related to the math-ability of these students. Others will guess the right answer one-tenth of the time. In such case, the question again will appear as if it is highly *specific*, but the reason for specificity is due the chance.

This empirical specification is very similar to the latent trait estimation methodology developed in the psychometrics field to design and evaluate tests called "Item Response Theory" (IRT) (Lord, 2012). This statistical technique was developed to measure latent traits associated with individuals and test items.[6] With IRT, item difficulties and discrimination levels are obtained along with the individual abilities. The latent structure analysis started with Lazarsfeld (1950), and the probabilistic item factor model was developed by Lord (1952). Rasch (1960) was first to use logit functions in predicting the latent traits and Birnbaum (1968) improved the methodology to allow different item difficulty levels to be predicted.[7]

To illustrate how we can use our framework to estimate production ability and product sophistication we will perform two applications. We first use international trade data, to measure the productive ability of countries and the sophistication of tradable products. Then we perform a second application using US city level employment using data, which allows to capture the sophistication of services and non-tradable industries.

When we apply our method to international trade data, we find that production ability have a strong positive correlation with GDP per capita levels. Interestingly, the main outliers are –as expected– countries that export mostly natural resource rich products. These countries are relatively rich but for reasons other than having a high production ability. When we turn to the estimates for products, we find that product sensitivity levels associated with natural resource products are often very low, meaning that there is not much selection on the ability of the countries that make them. The country production ability also reflects the potential of economic growth of countries. We estimate growth regressions and find that a country's ability, controlling for the current level of income

[6]The framework we propose is similar to what is known as the 2 parameter (2-pl) model in IRT. What its call in IRT item discrimination parameter is what we call in this paper product sensitivity, and item difficulty is what we call product specificity.

[7]We take advantage of the well-developed empirical methods of IRT to estimate the product sensitivity and specificity parameters as well as the location abilities.

per capita, is associated with a larger growth rate, a finding that remains even after controlling for country fixed effects.

Our estimates on production ability could be picking up other specific features of countries important for competitive advantage, such as a sound monetary policy, a functioning rule of law and contract enforcement. So we turn to a second application using subnational data, where the country level features are constant across regions and the ability levels of the regions will determine the industrial structure.. Specifically we use industry-employment data at the metropolitan statistical area (MSA)[8] level of the United States. We confirm the previous results using trade data by finding that cities with high ability have higher income per capita, and that production ability has also a strong correlation with growth of income and employment growth. Also in line with the result when using trade data, we find that industries reliant on minerals or extensive land, tend to have high levels of specificity. And manufacturing of machinery, and high skilled services are among the industries with higher sensitivity level towards production ability.

The rest of the paper is structured as follows. In the next section, we will conduct a literature review, linking this paper to previous contributions. Then, in Section 3 we introduce the capability-based model and explain how we can obtain product sophistication and country ability measures. Section 4 will describe the datasets and the empirical strategies that we use in the rest of the paper. Section 5 will report our results in both international and subnational level. Finally, we will conclude in Section 6.

2 Literature review

There are some influential papers proposing measures that summarize the production composition of countries and cities. For instance Glaeser et al. (1992) studies the employment pattern of industries and cities, and show that the diversity of industries is a strong predictor of city growth. At the country level, Hausmann et al. (2007) propose measures that summarize the income level of export basket and products, showing that what countries export is informative of the their prosperity paths. Hidalgo and Hausmann (2009) follow the spirit of two papers already mentioned, but proposes measures for countries and products that just focus on the diversity of products and their ubiquity.

[8]Throughout the paper we use city interchangeably with MSA)

Overall, this kind of measures have gained popularity recently, becoming important complements to other measures as tools for policymaking (Love and Stockdale-Otárola, eds, 2017).[9]

Our work is intimately related to the Economic Complexity Index (ECI), which was first algorithmically introduced in Hidalgo and Hausmann (2009), and later formalized in Hausmann et al. (2014). Importantly ECI has a high positive correlation with GDP per capita, and is a strong predictor of countries' economic growth. The algorithm behind ECI attempts to capture the capabilities of countries and products by using product ubiquity to correct for the diversity of countries. But there is not a straight mapping between the capabilities and the complexity measures. In our case, however, the country ability measure is related to the norm of the capability vector of the countries. Moreover, the original interpretation of ECI has come under scrutiny by Kemp-Benedict (2014), which suggests that ECI carries information orthogonal to the diversity for countries contradicting the original motivation behind this measure. An alternative interpretation of ECI is given by Mealy et al. (2019), where the authors argue that ECI algorithm is, in a sense, a dimensionality reduction method, equivalent to spectral clustering algorithms used in partitioning similarity graphs. Yet, this interpretation is not supported with an underlying economic model. Our methodology here has a clear interpretation, with the measures directly associated with the capability vectors of the countries and products. In terms of economic importance, when we compare our production ability measure of countries and cities to ECI in predicting economic growth, the ability turns out to be a better predictor of future economic growth.

The idea to reward diversity of countries but to penalize for ubiquity of products gave rise to other measures such as those proposed by Tacchella et al. (2012), labeled by the authors as country fitness and product quality. There has been, however, some questions about the convergence of this algorithm (Morrison et al., 2017) and the dynamics associated with this algorithm suggest heterogeneous dynamics in association with economic growth (Cristelli et al., 2015). In our analysis we also include this measures, and find that our variables of product sophistication and production ability measures have a more robust correlation with income levels and economic growth.

Aside from the relationship between variables that summarize the production capability of loca-

[9]There is another strand of literature developed or surveyed by Brahmakulam et al. (2001); Desai et al. (2002); Lall (2003); Archibugi and Coco (2005) that focuses on indexing technological capabilities using patents, R&D spending, scientific publications, and physical and social infrastructure, etc. that is beyond the scope of this paper. Here, we focus on the productive ability inferred indirectly from the patterns of industries or technologies.

tions and income levels, Hartmann et al. (2017) links economic complexity to the inequality levels observed across countries. The product complexity measures, on the other hand, have been used as an indicator of product desirability. For instance, Javorcik et al. (2018) shows that foreign direct investment results in an increase in the product complexity levels of regions in Turkey. Felipe et al. (2012a) characterizes the product complexity index and Felipe et al. (2012b) analyzes China's development in light of increasing product complexity. Hausmann et al. (2014) introduce a measure called Complexity Outlook Index (COI), which complements ECI in determining a country's growth potential by analyzing how well a country is situated vis-à-vis complex products in the product space. Product sophistication can be also viewed as the step size while climbing the quality ladders (Verhoogen, 2008; Hallak, 2010; Khandelwal, 2010) or technological development (Petralia et al., 2017). Overall, we believe that our product sophistication and country ability measures could be used in conjunction with already established measures to find the determinants of economic growth and structural change (Teixeira and Queirós, 2016).

Economic complexity measures were originally developed for the international trade data and have been extended to subnational data recently (Mealy et al., 2019; Fritz and Manduca, 2019). An economic complexity inspired algorithm applied to patent classification and US metropolitan areas shows that complexity affects the spatial distribution of knowledge (Balland and Rigby, 2017) and this approach supplements the previous measure of patent complexity proposed by Fleming and Sorenson (2001). Applied to European regions, the complexity of the patent classes imply that moving into higher complexity areas is beneficial but it is hard to achieve (Balland et al., 2018b). Finally, Balland et al. (2018a) analyze the urban concentration of complex economic activities and find that these types of activities have been increasingly concentrated in urban areas since 1850. Our sophistication and ability measure could be readily applied to determine innovational capacity of regions as well.

Our results are also related to diversification of regions and we find that our model captures the appearance of industries. In the diversification literature, the principle of related diversification (Hidalgo et al., 2018; Boschma, 2017) has been shown to influence the industrial patterns of regions. Relatedness could be though as a measure of shared capabilities between products or industries. In that regard, the regional diversification is linked with our production ability measure by default. Hidalgo et al. (2007) introduce the concept of product space, showing that the presence

of related products is an important determinants of the pattern of development of countries. Hausmann et al. (2019) further develops this idea and shows that the implied comparative advantage of locations could be inferred from the related industries, and this measure is highly correlated with the current revealed comparative advantage and predictive of the future comparative advantage. Boschma and Capone (2015) shows the importance of institutions in shaping the diversification using the measures from the product space. The complexity or ability/sophistication measures could give more directed targets to countries, so they might also play a role in diversification patterns of countries. Moreover, the industrial clusters have been found to cause differential regional growth (Porter, 2003) and firm performance (Lee, 2018). The regional diversification literature have been also debating whether specialization suggested by Marshall (1890) or diversification, emphasized in Jacobs (1969), have been more important in establishing sustainable cities (Beaudry and Schiffauerova, 2009). The diversification could also be related to the occupational structures of the cities which has been shown to affect their productivities (Bettencourt et al., 2014). Our model clearly favors the diversification process, which is consistent with the findings by (Glaeser et al., 1992), but our results still have room for improvements to include some sort of smart specialization patterns (Balland et al., 2018b).

3 Model

We will build on the model proposed by Hausmann and Hidalgo (2011). According to this model, to successfully make a product, locations need to have all necessary capabilities available, similar to the production function suggested by Leontief et al. (1953). Let's denote the capability requirement vector of product p with $k_p \in \{0,1\}^m$ where m is the number of capabilities, a vector length that in practice may be unknown. Location c will make a given product if and only if it has *all* the required capabilities. Formally:

$$M_{c,p} = \begin{cases} 1 & \text{if} \quad k_p \cdot k_p^{\mathsf{T}} = h_c \cdot k_p^{\mathsf{T}} \\ 0 & \text{otherwise} \end{cases}$$

where the dot operator $(\cdot)$ is the usual vector inner product operator, $\intercal$ is the transpose operator, $h_c \in \{0,1\}^m$ denotes the ability vector of the location l, and $M_{c,p}$ is the matrix whose entry (c,p) is equal to one if the location c makes the product p, and zero otherwise.

To give a simple example, let's assume there are 3 capabilities and a single product. The capability requirement vector for product p is $k_p = [1,1,0]$, meaning that the product requires the first and the second capabilities, but does require the third one. The squared norm of this vector is $k_p \cdot k_p^\intercal = 2$. Suppose country c has capabilities $h_c = [1,1,1]$. According to our model the country will be able to make that product because it has all necessary capabilities required. And mathematically we can confirm this since $h_c \cdot k_p^\intercal = 2$, which is equal to the squared norm of k_p. On the other hand, let's assume another country, c', with capabilities $h_{c'} = [1,0,1]$. Since $h_{c'} \cdot k_p^\intercal = 1 \neq k_p \cdot k_p^\intercal$, this country will not be able to make product p. The reason is that this country is missing the second capability, which is required by the product.

According to this model, the product capability space is multi-dimensional, which implies that by accumulating capabilities, countries can choose between multiple paths of development. But here, for empirical purposes, we will make a simplifying assumption and will focus on finding a single ladder of development. We will assume that the capability vectors show a nested structure. Nestedness here means that if we take any products, i and j, assuming that product i requires less capabilities than product j, then the capabilities required of product i will be a subset of capabilities required by product j. Hence, the capabilities required by product i could be found nested under the capabilities of product j for any (i,j) pair. Formally, we can index products by the norm of their capability vectors, with product p_1 having the smallest norm and product p_n having the largest norm. With the nestedness assumption, we can write:

$$k_{p_i} \cdot k_{p_i}^\intercal = k_{p_i} \cdot k_{p_j}^\intercal$$

for any $1 \leq i < j \leq n$. The nestedness property implies that we can keep track of products by their norms, which makes the problem at hand inherently one dimensional. In particular, we denote the norm of product p with $b_p \in \mathbb{R}^+$m which captures the extent of capabilities required by each product.

If we assume that product capability vectors are nested, a country that makes product i will

jump to the $i+1$th product by acquiring the missing capabilities required as if climbing up a ladder formed by products. Hence, we can match a country's ability level to a particular product sophistication level by finding the most sophisticated product the country makes. Formally, let's denote the norm of a country capability vector with θ_c. Country c will make product p if $\theta_c \geq b_p$. In the nested case, this implies that the country c has all necessary capabilities to make product p. In other words, the probability of a country making a product will be a step function[10] In this scenario, θ_c becomes a threshold level for the country; any product whose sophistication is below or equal to this level will be made by the country. Any product with higher sophistication than this level will not be made the country.

This process, in reality, may be more relaxed than the stringent combination of capabilities implied by the Leontieff production function. First, there could be some tolerance to missing capabilities, or at least some degree of elasticity of substitution between capabilities. To allow for substitution between capabilities, we make the probability of making product p in country c an increasing function of $(\theta_c - b_p)$. In that case, as θ_c increases and approaches the b_p level, the likelihood of making a product will also increase. Once θ_c surpasses the product sophistication level, as the difference increase, the probability of making the product will approach zero.

Second, we make the sensitivity paramter different across products. This assumption is also realistic, since some products –like copper or oil– do not depend highly of other production factors –such as skilled labor–, but rather their presence in a city or country is a result of geological lottery. Hence, these products would not be sensitive towards the heterogeneity of capability across places. In particular, we can introduce the (discrimination) parameter, a_p and make the probability of country c producing product p as a function of $a_p\theta_c - b_p$. In that case, depending on a_p, the probability of making the product will be different.

Third, we would like to choose a functional form of probabilities such that as a_p becomes larger likelihood becomes very similar to the Leontief model. Moreover, the functional form should satisfy the property that as $(a_p\theta_c - b_p) \to -\infty$ the limit probability will be zero, and as $(a_p\theta_c - b_p) \to \infty$ the limit probability will be one. A probability function that satisfies these properties is the logistic function. In particular, we can write the probability of location c making product p as a function of

[10] See figure 1 below. In the figure of the right panel we show the Leontieff case depicted with the black line.

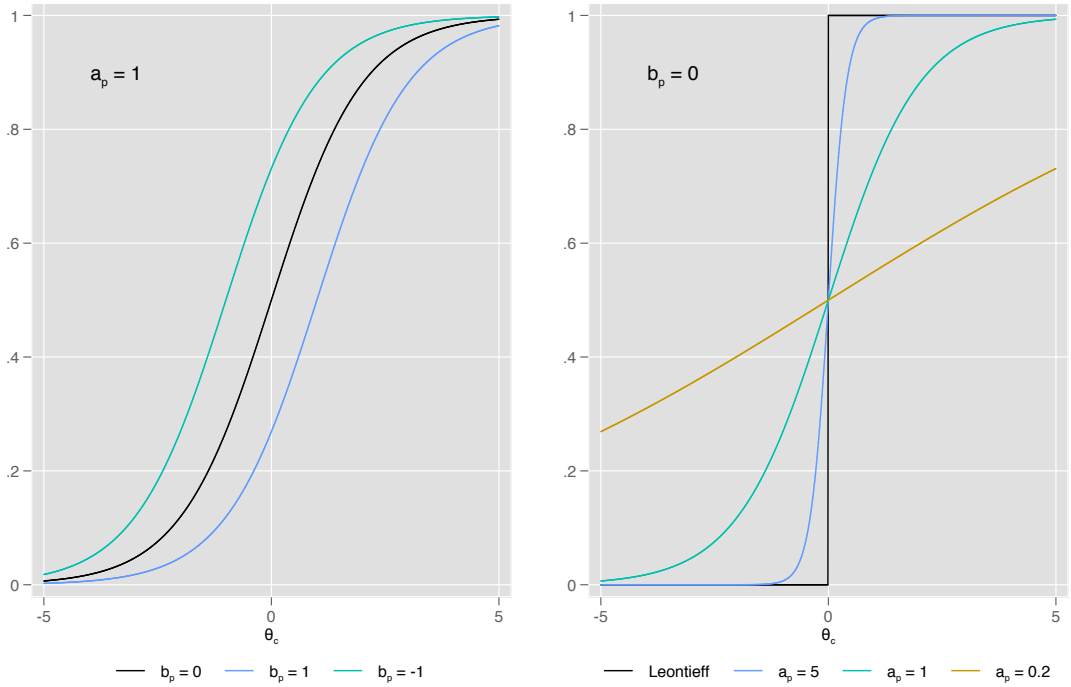
country ability (θ_c) as:

$$P(M_{c,p} = 1|\theta_c, a_p, b_p) = \frac{1}{1 + e^{-(a_p\theta_c - b_p)}}. \quad (1)$$

The left panel of figure 1 shows the logistic function for different values of b_p, which captures the *product specificity*. As b_p changes, the probability curves shift in the real axis and as b_p grows, fewer locations are likely to make the product. The right panel shows the logistic function for different values of a_p, which captures the *product sensitivity*. As a_p increases, the functional form better captures the step function implied by the Leontieff production function, which corresponds to the black line in the figure. As a_p decreases the the choice of θ_c has less of an impact on the likelihood of a location making the product.

Interestingly, a product with high b_p and very low a_p will be likely made by few locations, and the list will not be explained heavily on the production ability (θ_c) of that location. Likewise, a product with both high b_p and a_p will be made by few locations as well, and the set of producers will depend on their ability. Examples for the first set of products can be found in the natural resource products whereas the second set of products are mostly knowledge-intensive products.

Figure 1: Probability distributions



In our empirical estimations, we use the information about which location makes a given product to predict the sophistications of product and abilities of countries. Using the logistic probability model we introduced, we can write the likelihood of observing the location-product pairing matrix, $M_{c,p}$ as:

$$L(M|\Theta, B, A) = \prod_{(c,p) \text{ s.t. } M_{c,p}=1} \frac{1}{1+e^{-(a_p\theta_c - b_p)}} \prod_{(c,p) \text{ s.t. } M_{c,p}=0} \left(\frac{e^{-(a_p\theta_c - b_p)}}{1+e^{-(a_p\theta_c - b_p)}} \right) \tag{2}$$

where Θ denotes the set of location production ability (θ_c's), A denotes the set of product sensitivity levels (a_p's) and B denotes the set product specificity levels (b_p's). The parameters $\hat{\theta}_c$, $\hat{a}_p$ and $\hat{b}_p$ are those that maximize the likelihood function.

Coincidentally, as mentioned above, the maximum likelihood problem we posed here is similar to the well-known psychometric method called the Item Response Theory (IRT), which is widely used to design and assess tests or questionnaires. In the IRT formulation, each individual taking the test has an ability and each question has a discrimination and a difficulty level. The probability of answering a question right depends on the difference between the ability of the individual and the difficulty of the question. In our case, the questions are analogous with the products and individuals are analogous with the locations. IRT also employs the logistic functional forms as probability density functions so we can use estimation the techniques developed in that field for the problem in hand.

4 Empirics

We apply our methodology to countries and cities, using international trade and city-industry data. For international trade, we use the cleaned version of UN COMTRADE international trade data by Bustos and Yildirim (2019). Specifically, we use the country-commodity exports data following the SITC Rev. 2 classification between years 1962 and 2016. We follow the list of countries used by Hausmann et al. (2014), eliminating countries whose exports are below $1 billion in 2008, and those that experience war, or have severe trade data problems leading to unreliable reports of the trade. This leaves us with a list of 124 countries. For products, we eliminate those ones with small export volumes (summing up to 0.5% of world trade) and those that we do not observe reported throughout the sample duration, giving us 660 products. We used the World Development Indicators (WDI)

for the country level income and population data.

For the subnational results, we use the County Business Patterns (CBP) database from 2005 to 2015 by the US Census Bureau. We convert all the data to follow the NAICS 2002 industrial classification at 4 digits, using the concordance tables provided by the Census Bureau. As a geographical unit, we use metropolitan statistical areas (MSA) resulting in 764 MSAs and 266 industries. We obtained MSA-level income per capita data, and Input-Output tables from the U.S. Bureau of Economic Analysis (BEA). We also use the Occupational Employment Statistics (OES) data from the Bureau of Labor Statistics (BLS), and use the O*NET data to calculate the average training intensity of industries.

Our method uses a binarized version of the industry level data, both for the international trade data and subnational employment data. For the international data, we use the Revealed Comparative Advantage (RCA) measure first introduced by Balassa (1965) and from that measure we develop the binary country-product matrix, M following Hidalgo et al. (2007). Formally, let's assume that the $X_{c,p}$ is equal to the total exports of country c in product p. We can define RCA as:

$$\text{RCA}_{c,p} = \frac{X_{c,p} / \sum_c X_{c,p}}{\sum_p X_{c,p} / \sum_p \sum_c X_{c,p}} \tag{3}$$

The numerator of the RCA formula describes the share of country c in total exports of product p in the world, whereas the denominator is the share of country c in total world trade in all products. If the RCA value is higher than 1, it implies that the country c exports product p more than its fair share, which is captured by the denominator. Using this idea, we can obtain M from RCA matrix:

$$M_{c,p} = \begin{cases} 1 & \text{if RCA}_{c,p} \geq 1 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

For the subnational case, we will use the Location Quotient (LQ), which is equivalent to RCA but it using employment instead of the trade data. Suppose $\text{emp}_{c,p}$ denotes the employment in MSA or province c in industry p. Then, the LQ can be obtained replacing $X_{c,p}$ with $\text{emp}_{c,p}$ in Equation 3. For

LQ, the numerator is the employment share of industry i in location c, and the denominator is the share of location c in the total employment nationally. Similarly, M matrix for the subnational case can be defined using LQ instead of RCA in Equation 4.

4.1 Hypotheses

We expect the production ability parameter (θ_c) to capture the productive capability of a location, and the sensitivity and specificity of products to be related to other product quality measures. We also anticipate these parameters to shape the economic landscape of both location and industries. In particular:

H1: The country ability parameter will be highly correlated with ECI (Hausmann et al., 2014), Fitness measure of Tacchella et al. (2012), which use the same underlying data.
H1': At subnational level, we expect these relations to hold at the MSA level as well.

H2: The country ability measure will be correlated with the GDP per capita of the country. Furthermore, after controlling for the country's current income level and other country level measures, we expect the ability to be a significant predictor of economic growth measured as the growth in income per capita. Formally, if we estimate the following equation:

$$\text{Growth}_c = \beta_y \times lny_c + \beta_a \times \text{Ability}_c + \beta_x X_c + \epsilon_c \tag{5}$$

we expect to have $\beta_a > 0$ and $\beta_y < 0$. Here, lny_c is the logarithm of income per capita of country c and X_c vector represents the other country-level controls.
H2': At subnational level, we expect the MSA ability to be a significant predictor of MSA level income per capita growth.

H3: Product sensitivity and specificity measures will be highly correlated with PCI (Hausmann et al., 2014), and the Product Quality based on Fitness (PQF) measure of Tacchella et al. (2012). We expect sensitivity to be more correlated with PCI because they are both related to the product's overall selectivity whereas specificity to be more correlated with PQF measure since they are both

intimately related to the ubiquity of the products.
H3': At subnational level, we expect similar relationships between industry sensitivity and specificity measure with PCI and PQF measure to hold.

H4: A country becomes a natural resource exporter not based on its abilities but due to geographical luck. Moreover, adverse effect of natural resource exports on the productive structure of the countries, known as Dutch disease, has been well-documented (Corden, 1984). Therefore, we expect the ability levels of primary natural resource countries to be not significantly correlated with their income levels. And the corresponding sensitivity parameters of the natural resource exports to be low, whereas their specificity levels to be high. For some products like oil and natural gas, we might even see negative sensitivity parameters, implying that selection on ability might be reversed.

H5: At the subnational level, we expect industries that are present or close to be present in all cities such as auto-repair services to have both low specificity and low sensitivity.

5 Results

5.1 Country ability and product sophistication using international trade data

We estimate country production ability and product sophistication levels using the international trade data for each year separately, from 1962 to 2016.[11] Table 1 shows the top and bottom 10 countries according to our measure of production ability using data for year 2015. For comparison purposes we show the ability measure normalized to have mean of zero and a standard deviation of one. Germany is at the top, followed by Italy, Austria and the United States. China, a recent manufacturing powerhouse with a spectacular export performance in recent years is ranked at number 7. The table also shows the worst performers, including Nigeria, Libya and Venezuela. Interestingly, all countries in the the bottom performers are countries with extractive economies heavily reliant on natural resources, a list that includes Qatar and Kuwait –which have high levels of income per capita– but relying on a single extractive industry (oil and natural gas, respectively).

[11] We performed the IRT analysis using the Test analysis modules (TAM) package of R software (Robitzsch et al., 2018) to obtain estimates of location abilities, product sensitivity and product specificity levels. TAM package uses maximum likelihood estimation to calculate these estimates.

Table 1: List of Countries Ranked by Production Ability

Top countries

Country	isocode	ranking	ability
Germany	DEU	1	1.303
Italy	ITA	2	1.209
Austria	AUT	3	1.177
United States	USA	4	1.112
Czech Republic	CZE	5	1.066
Japan	JPN	6	1.053
China	CHN	7	1.023
Sweden	SWE	8	1.012
Spain	ESP	9	0.994
Slovenia	SVN	10	0.989

Bottom countries

Country	isocode	ranking	ability
Qatar	QAT	116	-1.722
Congo, Rep.	COG	117	-2.310
Kuwait	KWT	118	-2.311
Papua New Guinea	PNG	119	-2.468
Gabon	GAB	120	-2.469
Angola	AGO	121	-2.596
Algeria	DZA	122	-2.610
Venezuela, RB	VEN	123	-2.717
Libya	LBY	124	-2.776
Nigeria	NGA	125	-3.931

The left panel of Figure 2 shows the correlation between income levels of countries and the corresponding ability levels for 2015. Countries whose incomes rely heavily on natural resource products (e.g. Qatar, Kuwait, Norway, Gabon and Saudi Arabia) are outliers and appear above the regression line. This implies that, according to our theory, these are countries richer than what we expect given their production ability. In the right panel, we control for the natural resource exports intensity, finding that the (partial) correlation between ability and income increases from 39% to 68%, confirming our hypothesis about the natural resource exporters.

Figure 2: Country production ability and income per capita

(a) correlation



(b) controlling for natural resource exports



Aside from income levels, we also tested whether our ability measure is correlated with the other measures that summarize the composition of exports. In particular, we focused on the exports Diversity (Hidalgo and Hausmann, 2009), ECI (Hausmann et al., 2014) and country Fitness (Tacchella et al., 2012). Table 2 shows the correlation between these measures for year 2015. All these measures

are highly correlated with each other, showing that they are capturing common patterns from what countries export competitively to other countries. In figure A.1a we show the relationship between ECI and Ability, and the relationship appears to be non-linear. Once we compare the rankings of countries, the correlation between ECI rankings and Ability rankings rise up to 91% (Figure A.1b). Figure A.1c shows that in the period between 1962 and 2016, the correlation between ECI and the Ability is always high and fluctuates between 80-90%. But the correlation between the Ability and the per capita income has declined between those periods. The reason behind this decline can be attributed to the effect of commodity and natural resource products described above.

Table 2: Correlation between country measures

	Ability	ECI	Diversity	Fitness
Ability	1			
ECI	0.808***	1		
Diversity	0.820***	0.804***	1	
Fitness	0.681***	0.872***	0.898***	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3 shows the top and bottom 10 products sorted by an overall ranking calculated by re-ranking the products based on their total ranks when ranked by product specificity and sensitivity.[12] The top products are mostly complex machinery, with nuts and bolts the only product outside this category[13] The ranking at the bottom is dominated by products associated mostly with the abundance of natural resources (such as oil) or products that require specific climate or soil to be competitive in their production.

[12]We also show the rankings separately. Table A.1 and A.2 in the appendix presents the top and bottom products according the Sensitivity and Specificity parameter estimates, respectively. In these tables, we observe that the sensitivity parameter severely punishes the natural resource products as we expected.

[13]Despite what could sound as simple product, the international trade of Nails nuts and bolts is dominated by high productivity countries. China, Germany, Taiwan, Japan and the US account for almost 2/3 of world trade in this product.

Table 3: List of Products

Top products

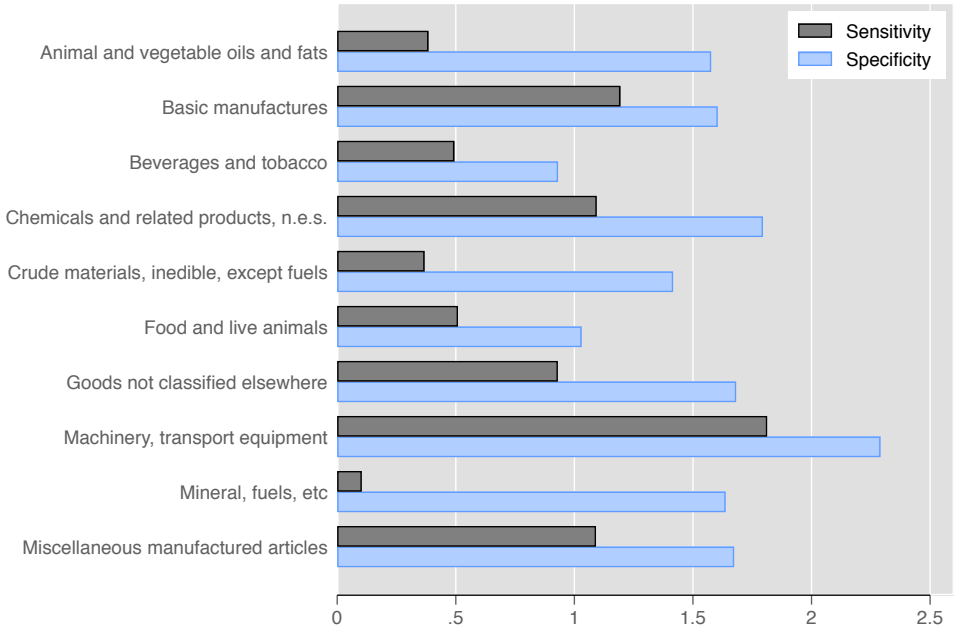
Products	SITC-4 Code	Ranking Overall	Ranking Specificity	Ranking Sensitivity
Nails, Nuts & Bolts	6940	1	3	1
Working Metal & Metal Carbides Machines N.E.S.	7367	2	1	3
Mechanical Tools For Building	7493	3	12	5
Rolling Mills	7372	4	7	11
Parts Of Paper Making Machines	7259	5	11	8
Internal Combustion Engines For Motor Vehicles	7132	6	13	6
Textile Machinery	7244	7	2	17
Dishwashers	7753	8	8	12
Reciprocating Pumps	7421	9	5	18
Valves	7492	10	24	7

Bottom products

Lubricating Petroleum Oils N.E.S.	3345	651	591	627
Frozen Fish, Excluding Fillets	0342	652	623	604
Fresh Or Dried Fruit N.E.S.	0579	653	653	574
Bovine & Equine Leather	6114	654	643	584
Wheat Or Meslin Meal Or Flour	0460	655	651	581
Sawlogs & Veneer Logs Of Non-Coniferous	2472	656	600	639
Worked Wood Of Non-Coniferous	2483	657	629	616
Bran, Sharps & Other Cereal Residues	0812	658	613	634
Unwrought Lead & Alloys	6851	659	641	628
Flora In Pharmacy	2924	660	645	638

Figure 3 shows averages for sensitivity and specificity parameters within one-digit SITC codes. As expected, the minerals and fuels section has the lowest average sensitivity, followed by the crude materials section. Nevertheless, these product sections with low-sensitivity do not exhibit relatively low average specificity compared to other sections. The machinery and transport section, on the other hand, has both the highest average sensitivity value and highest average specificity value. Other manufactured section follows the machinery and transport section in terms of average sensitivity.

Figure 3: Average Sensitivity and Specificity of industry groups



As we did with the country ability measure, we can analyze how the sensitivity and specificity measures are related with the product complexity measures reported in the literature. We use PCI (Hausmann et al., 2014), Product Quality based on Fitness (PQF) (Tacchella et al., 2012) and Ubiquity (Hidalgo and Hausmann, 2009) measures to compare with our measures. Table 4 shows the correlation between these measures for year 2015. The sensitivity is highly correlated with PCI whereas the specificity is highly correlated with the ubiquity.

Table 4: Correlation between product measures

	Specificity	Sensitivity	PCI	Ubiquity	Quality
Specificity	1				
Sensitivity	0.559***	1			
PCI	0.607***	0.828***	1		
Ubiquity	-0.880***	-0.305***	-0.513***	1	
Quality	0.763***	0.516***	0.662***	-0.617***	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Overall, our country Ability measure and product sensitivity and specificity measures correlate well with the cross-sectional measures like income per capita of the country, ECI or the PCI. Next, we will test whether these measures are also important in the dynamics of the country growth and diversification.

5.2 Economic Growth of Countries

Does our estimate of production ability predict economic growth? We tackle this question estimating how predictive our estimate of country ability is with GDP growth per capita over a span of 5 and 10 years. For this we estimate simple growth regressions – a la Barro –, controlling for investment (as a proxy of savings) and population growth[14]. Table 5 shows our basic results, all estimations done by OLS. Columns 1 to 5 present estimates were the dependent variable of annualized growth rate over 5 years, while columns 6 to 10 does the same for the annualized growth rate over 10 years. To facilitate comparison across variables and columns, we have standardized all explanatory variables to have mean zero and standard deviation of one. In column 1 we show the baseline estimate, which suggest that there is (unconditional) economic convergence. In column 2 we control for production ability, which indicates that a one standard deviation increase in ability is associated with a 0.76 percentage points increase in income per capita per year. When we control for population growth and investment rates over GDP, the size of the coefficient of production ability reduces almost by half, and it is only significant at the 90% confidence level. In columns 4 and 5 we repeat the exercise of the previous two columns, but now adding country fixed effects which capture country unobserved characteristics invariant over time, finding that ability remains with sizeable association with economic growth but with less statistical significance over time. Columns 6 to 10 repeat the exercise commented so far, and the interpretation of the results is similar. One interesting finding shown in this table is that our estimate of production ability remains correlated with economic growth in the subsequent 5 or 10 years, despite adding country fixed effects to the estimation. In the extensive empirical literature studying determinants of economic growth, there are not many variables that remain significant after being subjected to this level of scrutiny involving country fixed effects.

[14]We restrict the set of control variables to savings rates (proxied by investment to GDP) and population growth, since these are commonly used by growth models building on Solow model.

Table 5: Country Economic Growth

	5 year growth rate					10 year growth rate				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Income per capita, log	-0.29**	-0.59***	-0.88***	-4.99***	-4.90***	-0.40***	-0.73***	-1.07***	-5.02***	-4.70***
	(0.14)	(0.15)	(0.17)	(1.61)	(1.80)	(0.14)	(0.15)	(0.16)	(1.33)	(1.51)
Ability		0.76***	0.30*	0.74**	0.63*		0.83***	0.38**	0.76***	0.64*
		(0.15)	(0.15)	(0.30)	(0.37)		(0.17)	(0.18)	(0.28)	(0.35)
Population growth, %			-0.69***		0.06			-0.80***		-0.53*
			(0.13)		(0.33)			(0.13)		(0.31)
Investment (%GDP)			0.09***		0.06**			0.04*		0.02
			(0.02)		(0.03)			(0.02)		(0.03)
Observations	961	961	769	960	769	468	468	421	464	417
R-squared	0.08	0.13	0.19	0.41	0.45	0.11	0.19	0.29	0.61	0.62
Method	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
year FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
country FE				yes	yes				yes	yes

Robust standard errors in parentheses clustered by country
*** p<0.01, ** p<0.05, * p<0.1

As shown in Table 2, our ability measure if highly correlated with the other country measures, namely Diversity ECI and Fitness. Here, we analyze the impact of these variable on economic growth and put them in a horse-race with our production ability measure to infer which variable better captures the growth potential of countries. Table 6 shows estimates of growth regressions using the Diversity, ECI and the Fitness. We first introduce variables one-by-one in columns 1 to 4, and then we re-estimate the regressions using the Ability and each explanatory variables together. As expected, Diversity, ECI and Fitness are strongly correlated with economic growth (columns 1 to 4), with coefficients implying a significant impact annual growth rates ranging from 0.4 (for Fitness) to 0.84 (for ECI). We then analyze the effect of these variables on economic growth in the presence of the Ability variable. For instance, when we control for ability and diversity in the same regression (column 5), the first variable remains with a coefficient similar to that found in column 1, while the latter becomes not statistically different from zero. The column also shows the p-value of a statistical test where the null hypothesis is that the coefficients estimated for two variables have the same size. The p-value of 0.2 indicates that we cannot reject the null hypothesis, although probably the result is driven by the large standard error associated with the diversity estimate. In columns 6 and 7 we repeat the estimates, but using ECI and Fitness. The results indicate that none of this variables remain significant when we also control for Ability, while the latter remains with a strong correla-

tion. The null hypothesis about similar coefficients is only rejected largely for the case of Fitness. As robustness check, table A.3 in the Appendix repeats the exercise but using 10 year annualized growth rates as dependent variable and the results confirm that the findings are robust to different time horizons.

Table 6: Economic growth, competing measures

	5 year growth						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Income per capita, log	-0.59***	-0.67***	-0.86***	-0.53***	-0.64***	-0.66***	-0.58***
	(0.15)	(0.17)	(0.20)	(0.19)	(0.17)	(0.20)	(0.17)
Ability	0.76***				0.66***	0.65***	0.77***
	(0.15)				(0.21)	(0.21)	(0.16)
Diversity		0.70***			0.15		
		(0.17)			(0.23)		
Economic Complexity			0.84***			0.16	
			(0.18)			(0.26)	
Fitness				0.40**			-0.03
				(0.16)			(0.13)
Observations	961	961	961	961	961	961	961
R-squared	0.13	0.11	0.12	0.09	0.13	0.13	0.13
Method	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Test eq. coef.					.212	.263	.001

Robust standard errors in parentheses clustered by country
*** p<0.01, ** p<0.05, * p<0.1

5.3 Countries' new products

Next, we test whether our model captures the dynamics in the extensive margin. In particular, we define our probability parameter to be:

$$\text{Probability}_{cp} = \frac{1}{1 + e^{-(a_p \theta_c - b_p)}}. \tag{6}$$

where θ_c is the country ability, a_p is the product sensitivity and b_p is the product specificity. We tested whether products that were absent in the initial year would appear in the final year if they have higher probability values in the initial year. In other words, this is to test whether the mistakes in the initial year are corrected in the final year. Table 7 shows the results for different definitions of presence and absence, and different time horizons. Our dependent variable is equal to 1 if the product was absent in the initial year but became present in the final year. If the product does not

appear even though it was absent in the initial year, then the dependent variable is 0. In all cases, we only consider the products that were absent in the initial year. To test whether the thresholds of calling a product present or absent have an effect on our predictions, we use different threshold pairs. In column 1 and 4, a product is considered to be present if its $RCA_{cp} \geq 1$, and absent if $RCA_{cp} < 1$. But this absence threshold could be too high, since calling a product absent if its RCA is 0.9 might be too harsh. For that reason, in rest of the columns we call a product absent if $RCA_{cp} \leq 0.1$.[15] In column 2 and 5, the rule for presence is $RCA_{cp} \geq 0.5$ whereas in columns 3 and 6, the rule for presence is $RCA_{cp} \geq 1$. In all regressions we use OLS method, with fixed effects for country-year and product-year. To put it in the perspective, we use one variable (Ability) for the countries, and two variables (sensitivity and specificity) for the products, and generate the probability measures. Even after controlling for the year-country and year-product fixed effects, the probability measure is a significant predictor of the product appearance.[16]

Table 7: New Product Appearance

	5 Year			10 Year		
	(1)	(2)	(3)	(4)	(5)	(6)
Present:	$RCA_{cp}^{t+5} \geq 1$	$RCA_{cp}^{t+5} \geq 0.5$	$RCA_{cp}^{t+5} \geq 1$	$RCA_{cp}^{t+10} \geq 1$	$RCA_{cp}^{t+10} \geq 0.5$	$RCA_{cp}^{t+10} \geq 1$
Absent:	$RCA_{cp}^{t} < 1$	$RCA_{cp}^{t} \leq 0.1$	$RCA_{cp}^{t} \leq 0.1$	$RCA_{cp}^{t} < 1$	$RCA_{cp}^{t} \leq 0.1$	$RCA_{cp}^{t} \leq 0.1$
Probability	0.003***	0.004***	0.003***	0.006***	0.007***	0.005***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
Revealed Comparative Advantage (RCA)	0.282***	0.504***	0.218***	0.253***	0.666***	0.307***
	(0.002)	(0.013)	(0.009)	(0.002)	(0.020)	(0.015)
Observations	456,204	292,314	292,314	262,421	169,794	169,794
R-squared	0.778	0.852	0.923	0.815	0.853	0.919

Robust standard errors in parentheses clustered by country and product
*** p<0.01, ** p<0.05, * p<0.1

5.4 City production ability and industry sophistication

Our results using international trade data show that country production ability and product sophistication are important variables towards understanding the income level and growth potential of

[15] In columns 1 and 4, we kept the original definition of presence and absence that we use throughout the paper using RCA as formalized in in Equation 4. In columns 2,3,5 and 6 we took a moving average of RCA for three years to make the variable more stable.

[16] We could not build a similar measure for ECI-PCI, Fitness-Quality or Diversity-Ubiquity pairs and test them here, since there is no model-guided way to combine these measures to predict product appearances.. The probability parameter that we built from the country ability and product sophistication measures are directly implied by the model.

countries. Here, we extend the results to the subnational level using the industrial level employment patterns in the United States at Metropolitan Statistical Area (MSA) level, or simply cities for brevity. First, we estimate the Ability measures associated with each city. Table 8 shows the top and bottom metropolitan areas in the US according to the city ability estimates using data for year 2015. The ranking has Chicago at the top, followed by Portland and Minneapolis. New York and San Francisco are also listed among the top cities. At the bottom 10, we find MSAs intensive on a few industries or production heavily reliant on natural resources.

Table 8: List of Metropolitan Areas

Top cities

Name MSA	MSA code	ranking	ability
Chicago-Naperville-Joliet, IL-IN-WI	16980	1	2.937
Portland-Vancouver-Beaverton, OR-WA	38900	2	2.439
Minneapolis-St. Paul-Bloomington, MN-WI	33460	3	2.400
Milwaukee-Waukesha-West Allis, WI	33340	4	2.348
New York-Newark-Edison, NY-NJ-PA	35620	5	2.301
Cleveland-Elyria-Mentor, OH	17460	6	2.270
San Francisco-Oakland-Fremont, CA	41860	7	2.259
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	37980	8	2.254
Cincinnati-Middletown, OH-KY-IN	17140	9	2.242
Salt Lake City, UT	41620	10	2.240

Bottom cities

Las Vegas, NM	29780	818	-2.185
Indianola, MS	26940	819	-2.186
Lexington, NE	30420	820	-2.196
Alice, TX	10860	821	-2.300
Lincoln, IL	30660	822	-2.303
Cleveland, MS	17380	823	-2.344
Grants, NM	24380	824	-2.608
Levelland, TX	30220	825	-2.785
Kennett, MO	28380	826	-2.937
Rio Grande City, TX	40100	827	-3.222

City production ability levels are also correlated with the other complexity measures like ECI, Fitness and Diversity (Table 9), exhibiting correlations larger than 80% with each measure. The correlation levels are higher compared to the international case but these variables capture different aspects of MSA level economic activities.

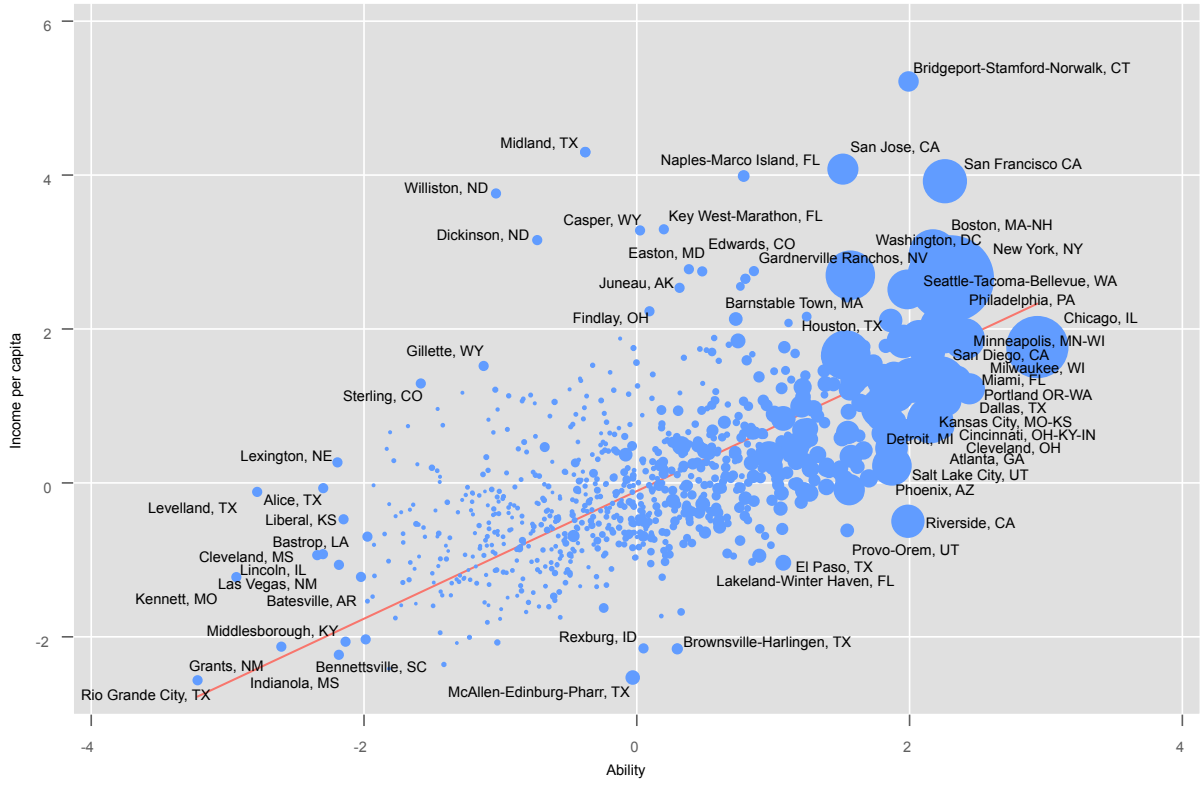
Table 9: Correlation between MSA measures

	Ability	ECI	Diversity	Fitness
Ability	1			
ECI	0.819***	1		
Diversity	0.925***	0.704***	1	
Fitness	0.845***	0.711***	0.849***	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

First, we explore whether the ability is correlated with the per capita income at the city level. Figure 4 shows the relation of average personal income and production ability using data for year 2015[17]. In the figure, bubbles are proportional to total number of workers in each city, according to the Census Bureau.[18] What we observe is that the both size and the income per capita are correlated with the ability levels of MSAs.

Figure 4: Production ability and average income in US cities



When we shift our attention to industries, the combined ranking of sensitivity and specificity

[17]Personal income data at MSA level was obtained from the Bureau of Economic Analysis.

[18]We exclude Jackson, WY from the figure since it is an outlier in terms of personal income: its income level is 8 times higher than the simple mean across MSAs.

puts the "Cable and Other Subscription Programming and Internet Publishing and Broadcasting" at the top (Table 10). Four out of top ten industries are coming from the Information sector. Some of the industries are entertainment and do not appear to be complex to produce, A possible interpretation of the industries in the list is that only complex cities are those in which we find significant presence of non-production industries, entertainment or amenities. The bottom of the list is populated with very ubiquitous industries such as Gasoline Stations with Convenience Stores, Restaurants, Religious Organizations and Nursing Care Facilities. The rankings does not change dramatically when we look at the rankings based on sensitivity and specificity separately as shown in the Appendix (Tables A.4 and A.5). Conspicuously missing among the top-ranked industries are manufacturing industries. This is consistent with the valuations of the firms in the stock markets as companies like Amazon, Microsoft and Facebook in the Information sector have higher values than the traditional manufacturing powerhouses like Ford and GM.

Table 10: List of Industries

Top Industries

Industry Title	NAICS-4 Code	Ranking Overall	Ranking Specificity	Ranking Sensitivity
Cable and Other Subscription Programming	5152	1	1	1
Internet Publishing and Broadcasting	5161	2	2	3
Software Publishers	5112	3	3	2
Investment Banking and Securities Dealing	5231	4	5	5
Pet Care (except Veterinary) Services	8129	5	8	4
Record Production	5122	6	4	8
Scheduled Passenger Air Transportation	4811	7	6	7
Miscellaneous Intermediation	5239	8	7	9
Theater Companies and Dinner Theaters	7111	9	13	10
Independent Artists, Writers, and Performers	7115	10	20	6

Bottom Industries

Used Household and Office Goods Moving	4842	257	250	250
Consumer Electronics and Appliances Rental	5322	258	257	254
Religious Organizations	8131	259	256	261
Limited-Service Restaurants	7222	260	262	256
Outdoor Power Equipment Stores	4442	261	261	259
Nursing Care Facilities	6231	262	258	263
Funeral Homes and Funeral Services	8122	263	260	262
Warehouse Clubs and Supercenters	4529	264	264	264
Automotive Parts and Accessories Stores	4413	265	265	265
Gasoline Stations with Convenience Stores	4471	266	266	266

The sensitivity and the specificity parameters also show high correlations with the PCI, Product

Quality based on Fitness (PQF) and Ubiquity measures, all estimated using data for year 2015 (table 11). Similar to the international case, the sensitivity is highly correlated with PCI, whereas the specificity is highly correlated with the ubiquity.

Table 11: Correlation between industry measures

	Specificity	Sensitivity	PCI	Ubiquity	Quality
Specificity	1				
Sensitivity	0.720***	1			
PCI	0.642***	0.913***	1		
Ubiquity	-0.923***	-0.587***	-0.441***	1	
Quality	0.768***	0.568***	0.640***	-0.507***	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The advantage of using the US subnational data is that we can analyze the relationship between the sensitivity and specificity parameters with many other industrial measures from various economic databases. Table 12 shows the results for different industry measures. By combining the O*NET's Education, Training, and Experience Category with the Occupational Employment Statistics (OES) data, we came up with the estimates of ratio of college graduates, average years of schooling in the industry, and average training required for the occupations in an industry. For all three measures, both sensitivity and specificity parameters are correlated. For the Years of Schooling and College Graduate Ratio variables, the correlation levels are not statistically significant but the specificity parameter exhibits a significantly higher correlation with the Training variable.

Next set of measures are calculates using the data coming from OES database. First of these measures is Occupational Herfindahl measure, which is the inverse Herfindahl of the occupational structure of the industries. Higher value in this measure implies higher number of effective occupations are present in the industry. The correlations show that the Specificity is correlated with the occupational Herfindahl measure, but not the sensitivity. The reason might stem from the manufacturing industries, which are specific but not at the top of the sensitivity list. Next measure is the logarithm of the employment in the industry, and this time the specificity measure is anti-correlated with the employment level. Last two measures from OES are the average hourly wage and the wage inequality measure calculated by dividing the hourly wage of a worker at 90th percentile with that of a worker at 10th percentile. Both measures are correlated with the sensitivity and the specificity.

The final three measures come from the Input-Output tables of year 2007. The Value Added Ratio is the ratio of value added part to total output of the industry. The sensitivity measure is cor-

related with the Value Added Ratio, but the specificity measure is weakly anti-correlated with it. Hence, industries with higher sensitivity parameters are proportionately contribute more to GDP. Input Herfindahl measure proxies the number of effective industries directly supplying to an industry and we do not find a clear correlation pattern here. The last measure is the total dollar amount required from all industries to produce one dollar of an output in an industry, called the Total Requirements. In this case, the specificity parameter is weakly positively correlated with this measure but the sensitivity parameter is negatively correlated with it. This effect might also be stemming from the manufacturing industries.

Table 12: Correlations with industry measures from real economy

	Sensitivity	Specificity
College Graduates, Ratio	0.436***	0.407***
	(0.0558)	(0.0570)
Years of Schooling	0.385***	0.361***
	(0.0597)	(0.0620)
Training	0.225***	0.394***
	(0.0674)	(0.0678)
Occupation Herfindahl	0.0499	0.261***
	(0.0570)	(0.0502)
Employment, Log	-0.102	-0.373***
	(0.0594)	(0.0566)
Avg. Hourly Wage	0.376***	0.513***
	(0.0622)	(0.0548)
Wage Ineq. (90/10)	0.379***	0.353***
	(0.0590)	(0.0609)
Value Added Ratio	0.197***	-0.114*
	(0.0543)	(0.0560)
Input Herfindahl	-0.0790	-0.108*
	(0.0524)	(0.0520)
Total Requirements	-0.250***	0.117*
	(0.0480)	(0.0504)
Observations	798	798

Robust standard errors in parentheses clustered by industry
*** p<0.01, ** p<0.05, * p<0.1

5.5 Growth of cities

Table 13 shows estimates of city growth over 5 and 10 years. We use two indicators of economic growth: growth of personal income, and growth of total city employment. Columns 1, 3, 5 and 7 present the baselines estimates. In each case it is interesting that there is convergence in income and

employment growth based on the initial levels of employment, at 5 and 10 year horizons, but income levels appear not to be meaningful as their statistical significance is well below the 90% confidence level. This result implies that small cities are places were income and employment is expected to grow faster than in large cities. In the other columns of the table we show estimates where we control for production ability, finding that it is strongly correlated with income and employment growth. The estimated coefficients imply that cities with an additional one standard deviation in ability are expected to have average income grow by 0.20%- 0.16% and employment grow by 0.48%-0.56% faster per year, according to estimates using 5 or 10 year horizon, respectively. Interestingly, the estimated coefficients are close to the size needed to overcome the convergence effect, and the estimate for employment growth is twice as large as that when the dependent variable is growth in income.

Table 13: City income and employment growth

	5 year				10 year			
	income		employment		income		employment	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Income per capita, logs	-0.04	-0.09	0.12	0.00	0.02	-0.02	0.07	-0.08
	(0.06)	(0.06)	(0.10)	(0.11)	(0.06)	(0.07)	(0.08)	(0.08)
Employment, logs	-0.11**	-0.24***	0.48***	0.17**	-0.14***	-0.24***	0.50***	0.16*
	(0.05)	(0.05)	(0.07)	(0.08)	(0.05)	(0.06)	(0.07)	(0.09)
Ability		0.21***		0.48***		0.16***		0.56***
		(0.04)		(0.09)		(0.04)		(0.08)
Observations	1,654	1,654	1,654	1,654	827	827	827	827
R-squared	0.02	0.03	0.21	0.22	0.02	0.04	0.11	0.15
Method	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
FE	year	year	year	year	year	year	year	year

Robust standard errors in parentheses clustered by city
*** p<0.01, ** p<0.05, * p<0.1

Are there other variables correlated with economic growth? As previously done for countries, we explore other proxies for the capabilities found in cities. Table 14 present the results, with income growth in columns 1 to 7 and employment growth in columns 8 to 14. As expected, ability, diversity, economic complexity and fitness are positively and statistically significantly correlated with income and employment growth. Interestingly, when we control for ability and use the other variables as additional controls, we find that for income growth only fitness survives as a control variable signif-

icantly different from zero. And for employment, only diversity remains as a control significantly different from zero. We repeat the test of equivalent coefficients when controlling for ability and the other 3 explanatory variables. The test delivers only two cases with p-value that is significant at conventional levels when comparing ability and economic complexity (see column 6 and 13). Overall, the picture that emerges is that despite the measure of economic growth used (either income or employment), or whether we make our estimate of production ability compete with other measures, productive ability remains strongly correlated and predictive of future economic growth.

Table 14: City income and employment growth under competing variables

	Income growth							Employment growth						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Income per capita, logs	-0.09	-0.05	-0.08	-0.09	-0.09	-0.09	-0.11*	0.00	0.10	0.01	0.05	0.02	0.00	-0.01
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.11)	(0.10)	(0.11)	(0.10)	(0.11)	(0.11)	(0.11)
Employment, logs	-0.24***	-0.22***	-0.15***	-0.32***	-0.27***	-0.25***	-0.35***	0.17**	0.18*	0.37***	0.14	0.07	0.17**	0.06
	(0.05)	(0.06)	(0.04)	(0.07)	(0.06)	(0.06)	(0.07)	(0.08)	(0.09)	(0.07)	(0.11)	(0.10)	(0.08)	(0.12)
Ability	0.21***				0.17***	0.25***	0.14***	0.48***				0.38***	0.49***	0.42***
	(0.04)				(0.05)	(0.08)	(0.05)	(0.09)				(0.09)	(0.14)	(0.10)
Diversity		0.15***			0.07				0.40***			0.23**		
		(0.05)			(0.05)				(0.10)			(0.11)		
Economic Complexity			0.11***			-0.05				0.30***			-0.00	
			(0.04)			(0.06)				(0.07)			(0.10)	
Fitness				0.28***			0.20***				0.43***			0.20
				(0.06)			(0.07)				(0.13)			(0.13)
Observations	1,654	1,654	1,654	1,654	1,654	1,654	1,654	1,654	1,654	1,654	1,654	1,654	1,654	1,654
R-squared	0.03	0.03	0.03	0.03	0.03	0.03	0.04	0.22	0.22	0.22	0.22	0.23	0.22	0.23
Method	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
FE	year	year	year	year	year	year	year	year	year	year	year	year	year	year
Test eq. coef.					.255	.025	.535					.377	.027	.233

Robust standard errors in parentheses clustered by city
*** p<0.01, ** p<0.05, * p<0.1

5.6 New Industries in Cities

We also tested whether the probability associated with our parameters defined in Equation 6 is predictive of industry appearances in cities. We use the same thresholds for the industry presences and absences as in the Table 7. Instead of RCA, we use the Location Quotient (LQ) measure to determine the presences and absences. Basically, we predict which industries will appear as present in 5 year or 10 year windows given the current level of probability. We predict the new industry appearances using OLS and controlling for MSA-year and industry-year fixed effects. Table 15 shows that higher probability estimated by our model in the initial year leads to a higher probability of appearance in the final year in all cases and time horizons. The coefficients associated with 5 and 10 year horizons

are close to each other. This result implies that our model is a significant predictor of diversification in the extensive margin.

Table 15: New Industry Appearance

	5 Year			10 Year		
	(1)	(2)	(3)	(4)	(5)	(6)
Present:	$LQ_{cp}^{t+5} \geq 1$	$LQ_{cp}^{t+5} \geq 0.5$	$LQ_{cp}^{t+5} \geq 1$	$LQ_{cp}^{t+10} \geq 1$	$LQ_{cp}^{t+10} \geq 0.5$	$LQ_{cp}^{t+10} \geq 1$
Absent:	$LQ_{cp}^{t} < 1$	$LQ_{cp}^{t} \leq 0.1$	$LQ_{cp}^{t} \leq 0.1$	$LQ_{cp}^{t} < 1$	$LQ_{cp}^{t} \leq 0.1$	$LQ_{cp}^{t} \leq 0.1$
Probability	0.028***	0.022***	0.015***	0.032***	0.025***	0.015***
	(0.001)	(0.003)	(0.002)	(0.001)	(0.003)	(0.002)
Location Quotient (LQ)	0.197***	-0.047	-0.011	0.134***	-0.102**	-0.067*
	(0.002)	(0.040)	(0.032)	(0.002)	(0.044)	(0.036)
Observations	343,557	51,444	51,444	229,832	51,444	51,444
R-squared	0.731	0.764	0.855	0.751	0.715	0.819

Robust standard errors in parentheses clustered by city and industry
*** p<0.01, ** p<0.05, * p<0.1

6 Conclusions

In this paper we have developed new measures that, we believe, capture the productive ability of countries and sophistication of products. Our measures are motivated through a simple capability based model, which can be estimated empirically using a maximum-likelihood approach.

We believe that there is a merit in doing the empirical analysis at both country and subnational levels. Countries not only differ by their ability levels but also their institutional quality might also play an important role for other abilities to be fully utilized. At the subnational level, variations associated with country properties like the institutions, monetary policy, rule of law will remain the same.

Our production ability captures information about the productive capability embedded in countries and cities. Just by using information on which products or industries are produced, we arrive to measures that strongly correlated with income levels and growth prospects. Production ability and sophistication of products introduces in this papers are correlated with the previously established measures of country complexity in the literature. But it also adds valuable information in predicting economic growth compared to the other measures.

Our method estimates two distinct measures for each product; sensitivity and specificity. These capture different aspects of product sophistication. The sensitivity parameter is highly correlated with PCI and the specificity is more tightly linked with the diversity. We also showed that these parameters are differentially correlated with the real economic parameters obtained from the input-output tables or occupational employment statistics. There are still more studies needed to understand the true nature of these parameters.

As an improvement compared to previous complexity measures like ECI or Fitness, our measures have narrower and clearer interpretation due to their a close link to a capability-based model of production. Algorithmic approaches minimize an underlying maximum-likelihood function or a free energy function (McEliece and Yildirim, 2003). But both for the complexity measures or fitness measure, the underlying functional forms were not formulated. Moreover, without knowing the functional form, we cannot associate any error terms with the measures. In the case of country ability and the product sophistication measures, however, we know the underlying maximum-likelihood function so from there we can calculate the standard errors of the estimated parameters.[19]

A simple extension of the estimation technique would be to add a baseline probability for each product and incorporate that into the maximum-likelihood function. With the baseline probability parameter, one can further discriminate between the products that do not require capabilities to be present and, rather their presence is due to geographical or other sort of luck. Another avenue could be to increase the dimensionality of the capability space and estimating the ability and sophistication parameters for more than one dimension.

[19] We did not use the error terms explicitly in our analysis in this paper, but they are available for the future use.

References

Archibugi, Daniele and Alberto Coco, "Measuring technological capabilities at the country level: A survey and a menu for choice," *Research Policy*, 2005, *34* (2), 175–194.

Balassa, Bela, "Trade liberalisation and "revealed" comparative advantage," *The manchester school*, 1965, *33* (2), 99–123.

Balland, Pierre-Alexandre and David Rigby, "The geography of complex knowledge," *Economic Geography*, 2017, *93* (1), 1–23.

_ , Cristian Jara-Figueroa, Sergio Petralia, Mathieu Steijn, David Rigby, and César Hidalgo, "Complex Economic Activities Concentrate in Large Cities," 2018. Utrecht University, Urban & Regional Research Centre, Papers in Evolutionary Economic Geography No: 18.29.

_ , Ron Boschma, Joan Crespo, and David L Rigby, "Smart specialization policy in the European Union: relatedness, knowledge complexity and regional diversification," *Regional Studies*, 2018, pp. 1–17.

Beaudry, Catherine and Andrea Schiffauerova, "Who's right, Marshall or Jacobs? The localization versus urbanization debate," *Research Policy*, 2009, *38* (2), 318–337.

Bettencourt, Luís MA, Horacio Samaniego, and Hyejin Youn, "Professional diversity and the productivity of cities," *Scientific Reports*, 2014, *4*, 5393.

Birnbaum, Allan, "Some Latent Trait Models and Their Use in Inferring an Examinee's Ability," in Frederic M. Lord and Melvin R. Novick, eds., *Statistical Theories of Mental Test Scores*, Reading, Massachusetts: Addison-Wesley, 1968, chapter 17-20, pp. 397–479.

Boschma, Ron, "Relatedness as driver of regional diversification: A research agenda," *Regional Studies*, 2017, *51* (3), 351–364.

_ and Gianluca Capone, "Institutions and diversification: Related versus unrelated diversification in a varieties of capitalism framework," *Research Policy*, 2015, *44* (10), 1902–1914.

Brahmakulam, Irene, Brian Jackson, Caroline Wagner, Anny Wong, and Tarsuro Yoda, "Science and technology collaboration: Building capacity in developing countries," 2001. RAND Science and Technology. Santa Monica, CA-2001.

Bustos, Sebastián and Muhammed A Yildirim, "Uncovering trade flows," 2019. Unpublished mimeo, available upon request.

Corden, Warner Max, "Booming sector and Dutch disease economics: survey and consolidation," *Oxford Economic Papers*, 1984, *36* (3), 359–380.

Cristelli, Matthieu, Andrea Tacchella, and Luciano Pietronero, "The heterogeneous dynamics of economic complexity," *PloS one*, 2015, *10* (2), e0117174.

Desai, Meghnad, Sakiko Fukuda-Parr, Claes Johansson, and Fransisco Sagasti, "Measuring the technology achievement of nations and the capacity to participate in the network age," *Journal of Human Development*, 2002, *3* (1), 95–122.

Felipe, Jesus, Utsav Kumar, Arnelyn Abdon, and Marife Bacate, "Product complexity and economic development," *Structural Change and Economic Dynamics*, 2012, *23* (1), 36–68.

_ , _ , **Norio Usui, and Arnelyn Abdon**, "Why has China succeeded? And why it will continue to do so," *Cambridge Journal of Economics*, 2012, *37* (4), 791–818.

Fleming, Lee and Olav Sorenson, "Technology as a complex adaptive system: evidence from patent data," *Research Policy*, 2001, *30* (7), 1019–1039.

Fritz, Benedikt SL and Robert A Manduca, "The Economic Complexity of US Metropolitan Areas," 2019. arXiv preprint arXiv:1901.08112.

Glaeser, Edward L, Hedi D Kallal, Jose A Scheinkman, and Andrei Shleifer, "Growth in cities," *Journal of Political Economy*, 1992, *100* (6), 1126–1152.

Hallak, Juan Carlos, "A product-quality view of the Linder hypothesis," *The Review of Economics and Statistics*, 2010, *92* (3), 453–466.

Hartmann, Dominik, Miguel R Guevara, Cristian Jara-Figueroa, Manuel Aristarán, and César A Hidalgo, "Linking economic complexity, institutions, and income inequality," *World Development*, 2017, *93*, 75–93.

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

_ , _ , **Daniel Stock, and Muhammed A Yildirim**, "Implied Comparative Advantage," 2019. CID Working Paper No. 276.

_ , _ , **Sebastián Bustos, Michele Coscia, Alexander Simoes, and Muhammed A. Yıldırım**, *The Atlas of Economic Complexity: Mapping Paths to Prosperity*, The MIT Press, 2014.

_ , **Jason Hwang, and Dani Rodrik**, "What you export matters," *Journal of Economic Growth*, 2007, *12* (1), 1–25.

Hidalgo, César A and Ricardo Hausmann, "The building blocks of economic complexity," *Proceedings of the National Academy of Sciences of the United States of America*, 2009, *106* (26), 10570–10575.

_ , **Bailey Klinger, A-L Barabási, and Ricardo Hausmann**, "The product space conditions the development of nations," *Science*, 2007, *317* (5837), 482–487.

_ , **Pierre-Alexandre Balland, Ron Boschma, Mercedes Delgado, Maryann Feldman, Koen Frenken, Edward Glaeser, Canfei He, Dieter F Kogler, Andrea Morrison et al.**, "The principle of relatedness," in "International Conference on Complex Systems" Springer 2018, pp. 451–457.

Jacobs, Jane, *The economy of cities*, Vintage, 1969.

Javorcik, Beata S., Alessia Lo Turco, and Daniela Maggioni, "New and Improved: Does FDI Boost Production Complexity in Host Countries?," *The Economic Journal*, 2018, *128* (614), 2507–2537.

Kemp-Benedict, Eric, "An interpretation and critique of the Method of Reflections," 2014. MPRA Paper No. 60705.

Khandelwal, Amit, "The long and short (of) quality ladders," *The Review of Economic Studies*, 2010, *77* (4), 1450–1476.

Lall, Sanjaya, "Indicators of the relative importance of IPRs in developing countries," *Research Policy*, 2003, *32* (9), 1657–1680.

Lazarsfeld, Paul Felix, "The logical and mathematical foundation of latent structure analysis," *Studies in Social Psychology in World War II Vol. IV: Measurement and Prediction*, 1950, pp. 362–412.

Lee, Chang-Yang, "Geographical clustering and firm growth: Differential growth performance among clustered firms," *Research Policy*, 2018, *47* (6), 1173–1184.

Leontief, Wassily et al., *Studies in the Structure of the American Economy*, Oxford University Press New York, 1953.

Lord, Frederic M., "A theory of test scores.," *Psychometrika Monographs. No. 7*, 1952.

_ , *Applications of item response theory to practical testing problems*, Routledge, 2012.

Love, Patrick and Julia Stockdale-Otárola, eds, *OECD Insights Debate the Issues: Complexity and Policy making* OECD Publishing Paris 2017.

Marshall, Alfred, *Principles of economics*, Macmillan and Company, 1890.

McEliece, Robert J and Muhammed Yildirim, "Belief propagation on partially ordered sets," in "Mathematical systems theory in biology, communications, computation, and finance," Springer, 2003, pp. 275–299.

Mealy, Penny, J. Doyne Farmer, and Alexander Teytelboym, "Interpreting economic complexity," *Science Advances*, 2019, *5* (1), eaau1705.

Morrison, Greg, Sergey V Buldyrev, Michele Imbruno, Omar Alonso Doria Arrieta, Armando Rungi, Massimo Riccaboni, and Fabio Pammolli, "On economic complexity and the fitness of nations," *Scientific Reports*, 2017, *7* (1), 15332.

Petralia, Sergio, Pierre-Alexandre Balland, and Andrea Morrison, "Climbing the ladder of technological development," *Research Policy*, 2017, *46* (5), 956–969.

Porter, Michael, "The economic performance of regions," *Regional Studies*, 2003, *37* (6-7), 545–546.

Rasch, Georg, *Studies in mathematical psychology: I. Probabilistic models for some intelligence and attainment tests.*, Nielsen & Lydiche, 1960.

Robitzsch, Alexander, Thomas Kiefer, and Margaret Wu, *TAM: Test analysis modules* 2018. R package version 3.0-21.

Tacchella, Andrea, Matthieu Cristelli, Guido Caldarelli, Andrea Gabrielli, and Luciano Pietronero, "A new metrics for countries' fitness and products' complexity," *Scientific Reports*, 2012, *2*, 723.

Teixeira, Aurora AC and Anabela SS Queirós, "Economic growth, human capital and structural change: A dynamic panel data analysis," *Research Policy*, 2016, *45* (8), 1636–1648.

Verhoogen, E.A., "Trade, quality upgrading, and wage inequality in the Mexican manufacturing sector," *The Quarterly Journal of Economics*, 2008, *123* (2), 489–530.

A Appendix

Additional Figures and Tables

Figure A.1: ECI and Ability

(a) score

(b) ranking



(c) Comparing correlation between ECI, Ability and GDP per capita over time

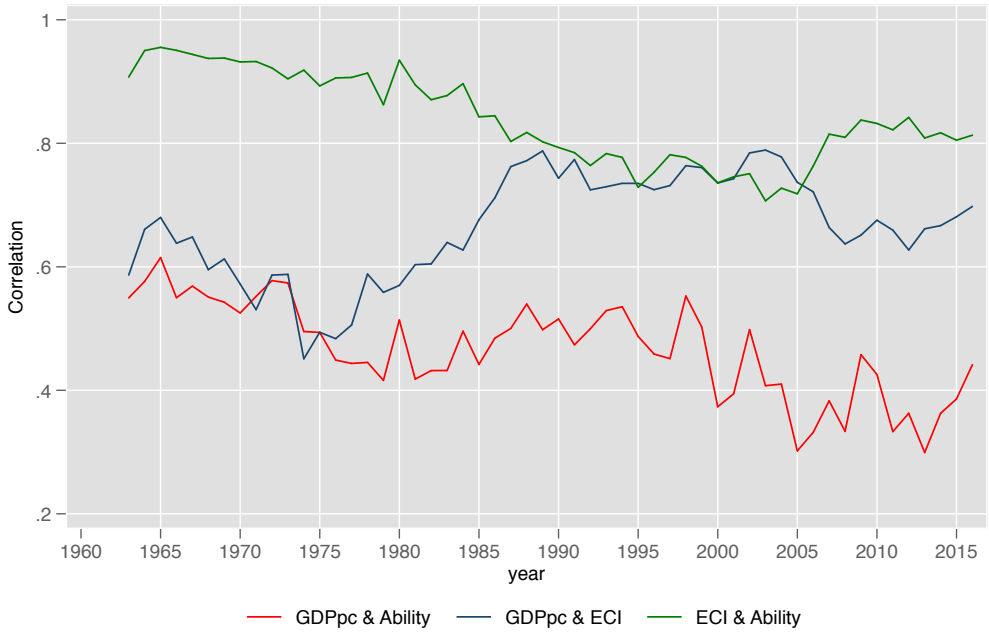


Table A.1: List of Products by Sensitivity

Top products

Products	SITC4 Code	Ranking	Sensitivity
Nails, Nuts & Bolts	6940	1	4.409
Other Vehicles Parts	7849	2	4.035
Working Metal & Metal Carbides Machines N.E.S.	7367	3	3.536
Articles Of Iron Or Steel N.E.S.	6997	4	3.488
Mechanical Tools For Building	7493	5	3.476
Internal Combustion Engines For Motor Vehicles	7132	6	3.476
Valves	7492	7	3.298
Parts Of Paper Making Machines	7259	8	3.066
Welding, Brazing & Cutting Machines & Appliances N.E.S.	7373	9	2.828
Other Articles Of Rubber	6289	10	2.803

Bottom products

Manganese	2877	651	-1.449
Other Non-Ferrous Base Metals	2879	652	-1.478
Not Agglomerated Iron Ore	2815	653	-1.542
Petroleum Gases	3414	654	-1.551
Other Sulphurs	2741	655	-1.555
Natural Rubber, Latex & Gums	2320	656	-1.565
Sesame Seeds	2225	657	-1.628
Raw & Roasted Cocoa Beans	0721	658	-1.679
Liquified Hydrocarbons	3413	659	-1.943
Crude Petroleum	3330	660	-2.382

Table A.2: List of Products by Specificity

Top products

Products	SITC4 Code	Ranking	Specificity
Working Metal & Metal Carbides Machines N.E.S.	7367	1	4.249
Textile Machinery	7244	2	3.905
Nails, Nuts & Bolts	6940	3	3.107
Wood-Based Panels	6344	4	3.015
Reciprocating Pumps	7421	5	2.839
Silicones	5827	6	2.604
Rolling Mills	7372	7	2.580
Dishwashers	7753	8	2.580
Asbestos	2784	9	2.528
Photographic Chemicals	8821	10	2.492

Bottom products

Wheat Or Meslin Meal Or Flour	0460	651	-1.846
Paper Packing Containers	6421	652	-1.914
Fresh Or Dried Fruit N.E.S.	0579	653	-1.983
Cement	6612	654	-2.029
Sugar Confectionary (Not Chocolate)	0620	655	-2.030
Iron & Steel Waste	2820	656	-2.206
Edible Products N.E.S.	0980	657	-2.223
Bakery	0484	658	-2.231
Closable Plastic Packing	8931	659	-2.416
Other Non-Ferrous Base Metals	2882	660	-2.856

Table A.3: Economic growth – 10 years – competing measures

	10 year growth						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Income per capita, log	-0.73***	-0.84***	-1.03***	-0.65***	-0.80***	-0.84***	-0.73***
	(0.15)	(0.17)	(0.20)	(0.20)	(0.17)	(0.20)	(0.17)
Ability	0.83***				0.67***	0.66***	0.83***
	(0.17)				(0.22)	(0.25)	(0.18)
Diversity		0.79***			0.24		
		(0.16)			(0.22)		
Economic Complexity			0.94***			0.26	
			(0.18)			(0.27)	
Fitness				0.44***			0.01
				(0.15)			(0.13)
Observations	468	468	468	468	468	468	468
R-squared	0.19	0.17	0.18	0.13	0.19	0.19	0.19
Method	OLS	OLS	OLS	OLS	OLS	OLS	OLS
year FE	yes	yes	yes	yes	yes	yes	yes
Test eq. coef.					.299	.402	.002

Robust standard errors in parentheses clustered by countries
*** p<0.01, ** p<0.05, * p<0.1

Table A.4: List of Industries by Sensitivity

Top Industries

Industry Title	NAICS-4	Ranking	Sensitivity
Cable and Other Subscription Programming	5152	1	3.960
Software Publishers	5112	2	2.931
Internet Publishing and Broadcasting	5161	3	2.690
Pet Care (except Veterinary) Services	8129	4	2.273
Investment Banking and Securities Dealing	5231	5	2.271
Independent Artists, Writers, and Performers	7115	6	2.242
Scheduled Passenger Air Transportation	4811	7	2.127
Record Production	5122	8	1.955
Miscellaneous Intermediation	5239	9	1.951
Theater Companies and Dinner Theaters	7111	10	1.881

Bottom Industries

Farm Machinery and Equipment Manufacturing	3331	257	-1.492
Carburetor, Piston, Piston Ring, and Valve Manufacturing	3363	258	-1.512
Outdoor Power Equipment Stores	4442	259	-1.646
Commercial and Industrial Machinery and Equipment (except Automotive and Electronic) Repair and Maintenance	8113	260	-1.681
Religious Organizations	8131	261	-1.704
Funeral Homes and Funeral Services	8122	262	-1.994
Nursing Care Facilities	6231	263	-2.133
Warehouse Clubs and Supercenters	4529	264	-3.160
Automotive Parts and Accessories Stores	4413	265	-3.218
Gasoline Stations with Convenience Stores	4471	266	-3.799

Table A.5: List of Industries by Specificity

Top Industries

Industry Title	NAICS-4	Ranking	Specificity
Cable and Other Subscription Programming	5152	1	5.029
Internet Publishing and Broadcasting	5161	2	2.890
Software Publishers	5112	3	2.736
Record Production	5122	4	2.669
Investment Banking and Securities Dealing	5231	5	2.521
Scheduled Passenger Air Transportation	4811	6	2.346
Miscellaneous Intermediation	5239	7	2.271
Pet Care (except Veterinary) Services	8129	8	1.903
Deep Sea Freight Transportation	4831	9	1.870
Iron Ore Mining	2122	10	1.687

Bottom Industries

Consumer Electronics and Appliances Rental	5322	257	-1.826
Nursing Care Facilities	6231	258	-1.864
New Car Dealers	4411	259	-1.872
Funeral Homes and Funeral Services	8122	260	-1.887
Outdoor Power Equipment Stores	4442	261	-2.134
Limited-Service Restaurants	7222	262	-2.157
Home Centers	4441	263	-2.204
Warehouse Clubs and Supercenters	4529	264	-2.997
Automotive Parts and Accessories Stores	4413	265	-3.355
Gasoline Stations with Convenience Stores	4471	266	-3.903