Industrial Structural Change and the Shifts in Comparative Advantage in Globalized Production

Ye, Xianjia*

August 25, 2016

Summary

Structural change in developing countries is usually understood as a shift of employment from agriculture to industry, and from less skill-intensive manufacturing (like textile) to more sophisticated and high-skilled ones (like machinery and electronics). With increasing globalization of production processes, this typology no longer holds true. Nowadays, most manufacturing products are produced in global value chains (GVCs), with different stages of production taking place in different regions in the world. This has major consequences for our understanding of development patterns (Taglioni and Winkler, 2016). In particular, it raises the need to analyse structural change as the shift in actual tasks carried out in an economy. Nowadays industrial upgrading is not merely the shift from one industry to another such as from textile to electronics, but also the shifts from less to more advanced tasks, such as from assembly tasks to R&D which take place within the same industry. As yet, our knowledge about this later type of structural change is very limited. This paper aims to contribute by using new data on tasks carried out in GVCs for a large set of products and countries. Its main finding is that in developing countries the structural change takes place most likely in a horizontal process in which labour shifts to another industry but perform tasks that require a same level of skill. Within industry upgrading is only rare; participation in GVCs is easy while the skill upgrading along the GVC is difficult.

More specifically, I follow Hidalgo et al (2007) and construct a *task space* built up from bilateral relatedness between tasks in various industries at different skill levels. This relatedness is calculated from the probability that a country has revealed comparative advantages (RCAs) in both tasks. Importantly, to control for offshoring I use value-added export to construct the RCA indices rather than the traditional gross exports. This is crucial as in GVC production gross export values contain a sizeable amount of foreign value added. Results show that a task is in general more related with other tasks at the same skill level, while the relatedness is low between different skilled tasks within an industry. For less developed countries it seems easy to participate in the low-skilled tasks of sophisticated global value chains (GVCs) and this is expected to bring a large gain. However vertical upgrading within industries to more skilled tasks is difficult to achieve. I also find that tasks in utility and logistics related services sectors have a strong complementarity with manufacturing tasks and they play an important role in structural upgrading. Probit regression shows that task relatedness has significant predicting power on actual directions of structural change.

*Ye: Faculty of Economics & Business, University of Groningen (email: x.ye@rug.nl).

1 Introduction

The export structure of a country is an important determinant of its economic growth. Lall *et al.* (2006), Hausmann and Rodrik (2003) and Hausmann, Hwang and Rodrik (2007) argue that products differ in their characteristics and their potentials to be a driver for future growth. The production of some products is more sophisticated as it entails more productive activities, uses better technology, and has larger network externalities and spillovers, such that these products provide more sources for future economic growth. This hypothesis is empirically confirmed in Hausmann, Hwang and Rodrik (2007) who show that the countries specialized in such growth-promoting products grow significantly faster.

While a good export structure predicts future growth, a more practical and policy related question remains untouched which is whether and how can a developing country upgrade towards a more favorable structure. For a long time, structural upgrading is understood as a horizontal process in which the employment and export of a country gradually shift across sectors, from agriculture and natural resource extraction to light industries like textile, and subsequently towards more modern manufacturing and the tertiary sectors (see, e.g. the famous Lewis model and Clark-Fisher model on structural change, Lewis 1954 and Clark 1940). However, the rise of offshoring in the recent decades has changed the natural of production and international trade. The production of final goods is now organized in the global value chains (GVCs), and the production process becomes increasingly fragmented across national borders. It seems to offer underdeveloped countries a shortcut in structural upgrading. Namely, an underdeveloped country first follows its comparative advantage and performs low-skilled tasks in the GVCs of sophisticated products, and then vertically climbs up the value chain through learning-by-doing and technical dissemination to obtain the comparative advantages in high-skilled tasks within the same industry (Lin and Wang 2008, Lin 2012, Taglioni and Winkler 2016).

Whether offshoring has indeed also changed the natural of structural upgrading? And what are the best practices for the policy makers to simulate the upgrading progress at different stages of development? To provide an answer to these questions, in this paper I try to empirically explore the potential structural upgrading paths when production is globally fragmented and offshoring is pervasive. In particular, I am interested in two types of upgrading, *horizontal* and *vertical*. The horizontal upgrading is the shift of employment *across industries* towards the tasks with a higher economic potential, say, re-allocating low-skilled workers from textile to the electronics assembly. The vertical upgrading refers to the process that a country vertically climbs up the value chain and upgrade towards higher-skilled tasks *within the same industry*, for example from low-skilled assembly tasks to high-skilled circuit engineering in the electronics industry. Using newly available data on task content of production and trade from 40 countries over the period of 1995 to 2009, I analyse the relatedness between each pair of task as proxied by the correlation in their revealed comparative advantages (RCAs) across countries. And based on these empirical relatedness indices I derive a network representation of production tasks that illustrate possible ways for upgrading. I find that the initial export structure of a country is closely related with the feasibility of the directions of its future structural change. And I will show that horizontal upgrading is relatively easy to achieve, while vertical upgrading is more difficult and gradual although it yields a higher gain in the long run. A different set of policy stimulations might be necessary to achieve vertical upgrading.

To illustrate the possible upgrading paths, I use a similar approach as the so-called "*product space*" framework in Hidalgo *et al.* (2007). The product space analyses the development path of a country's export structure by investigating the relatedness between products. The underlying notion is that the development of an economy is path-dependent, and the difficulty in entering the business of a new product depends on its relatedness with those products that the country is currently strong at. Similar as, say, a sneaker producer finds it easier to add badminton shoes into its product line rather than badminton rackets, at the macro level a new product is more likely to evolve in a country if it has strong tie with the country's current production and export structure. Hidalgo *et al.* empirically proxy the relatedness of a pair of products by the probability that both products have a revealed comparative advantage in a country, calculated on the basis of gross export data[1].

However, the upgrading in products (or industries) is an inappropriate perspective under globalized production. In the past when trade was only in final products, most tasks in producing a good were carried out within a country. In that case the exporter of a high-skill intensive good must perform many high-skilled tasks. But when the production process is globally fragmented, the products produced and exported by a country are no longer a good description for the actual production activities that are taking place. Offshoring brings the trade in tasks, and the comparative advantages of countries are now realized at task level, as shown in Baldwin (2006) and Grossman

1. The product space is essentially a "revealed relatedness" based on the observational co-existence of the RCAs of products, and it does not require ex-ante defined criteria for relatedness. Similar approaches can be found in Teece *et al.* (1994), Neffke and Henning (2008) and Bryce and Winter (2009). There are other measures that focus on specific and pre-defined aspects of relatedness, for example in Conley and Dupor (2003) and Neffke and Henning (2013). More discussions will follow in section 2.

and Rossi-Hansberg (2008). Empirical researches like Hummels *et al.* (2001), Becker *et al.* (2013) and Timmer *et al.* (2014) have also documented a rapid deepening in vertical specialisation and offshoring, and the increasing fragmentation of global value chains in the past decades. It is now not only feasible but also attractive for developing countries to participate in the GVCs of sophisticated and high-skill intensive products by performing low-skilled tasks. For example, Southeast Asia and China have witnessed a strong export growth in electronics since 1980s, but what they actually did was mainly assembly and this is far away from the "high-tech" stereotype of the electronics products they exported. On the other side of the coin, highly developed countries may also perform high-skilled tasks in the value chains of the so-called low-skilled products, like the design and global marketing of cloths, shoes and handbags and the organization of their material supply chain.

Therefore, focusing on the relatedness in products may give a misleading impression on the progress of structural change, and a task-based approach is needed which fits better the new production and trade paradigm with pervasive offshoring. In this paper I will follow Hidalgo *et al.*'s measure of relatedness, but I introduce the concept of *task space* that focuses on the relatedness between each pair of *tasks* instead of products, where a task is defined as a set of activities in an industry carried out by labour with a specific level of educational attainment (low-, medium-, and high-skilled). Furthermore, I will base my measures on the *value-added export* (VAE) of tasks, instead of gross export that includes the value of imported intermediates.

My task-based approach contributes to the current literature in three important aspects. First, it allows for the identification of the vertical alongside the horizontal types of structural change. The product space approach cannot identify the skill levels of tasks behind a product, and the current researches on the climbing-up-value-chain hypothesis are mainly from case studies of firms or specific industries. There is a lack of systemic research on the feasibility of vertical upgrading within different industries, and this paper aims to bridge this gap. By decomposing the activities of each industry into tasks at various skill levels, my task space approach is capable to evaluate the relative difficulties in both horizontal and vertical upgrading. As I will show later, the difficulties and benefits from two types of upgrading have large differences, therefore the identification between these two types is relevant.

Secondly, my task space approach uses value-added export (VAE) to calculate the RCAs instead of relying on gross export that is used by Hidalgo *et al.* VAE based RCAs take into account the complex structure of global production fragmentation, and subsequently give more reliable estimates for task relatedness. The gross-export based conventional RCA exaggerates the comparative advantage of countries located in downstream end of the GVC. This is because gross export of a product contains a considerable amount of imported intermediate inputs which cannot be distinguished from the value added by the country itself. As an extreme case, Dedrick *et al.* (2009) show that around 97% of the value in China's exports of iPod and HP laptops originated from outside China. The conventional RCAs is likely to over-estimate China's actual comparative advantages in electronics, and for countries specialized in upstream activities, like the US, underestimate. This bias is illustrated in Koopman *et al.* (2014) and Los *et al.* (2015) and is also confirmed in this paper. Using VAE resolves this problem since it excludes the value of intermediates and only focuses on the value-adding activities that take place within the country.

Thirdly, previous empirical structural change literature focus usually only on the manufacturing sectors. I also extend the scope of research to tasks in service sectors that become an increasingly important part of the global trade. I will include directly traded services, as well as indirect exported services from (non-traded) business supporting sectors that are embedded in the exported goods/services. This is not possible when gross export is used, but is possible in my approach since the measure of VAE can trace down to the exact (upstream) activities that generated the exported value-added. Therefore this approach allows me to analyze the important role of services in structural change.

The empirical contribution of this paper is to derive the "task space", i.e. a network illustration of potential structural upgrading paths in terms of production tasks based on tasks relatedness. Following Hidalgo *et al.* (2007), the relatedness between two tasks is defined as the probability that both tasks have a value-added based revealed comparative advantage in a country. For this, I first need a measure for VAE of each task, defined as the value-added created by a task that is ultimately absorbed in foreign final demand (Johnson and Noguera 2012). It will be derived from the World Input-Output Database (Timmer *et al.* 2015) which provides detailed cross country-industry supply/use tables and labour compensation per skill level. The WOID database covers 40 major economies in the world, including both advanced and developing countries. In total I include 28 sectors (ISIC 2-digit level), and I calculate the VAE and subsequently the RCA of a set of 84 unique tasks (each sector has activities at three skill levels[2]). The task space is then an 84×84 matrix; each element represents the bilateral relatedness between a pair of tasks, based on actual task level RCA indices from 40 countries.

2. A possible task can be, say, low-skilled activities in textile or high-skilled activities in electronics. Different skilled activities within an industry, for example low- and medium-skilled activities in electronics, are treated as different tasks.

The most important finding from my task space is that horizontal and vertical structural upgrading differ greatly in their propensity. I find that almost all low-skilled tasks in all industries are closely related with each other, meaning that it is relatively easy to re-allocate low-skilled workers horizontally across different industries. It points towards a strategy that in the early stage of development, it is feasible and worthwhile for a country to utilize their abundance of low-skilled labour in the low-skilled tasks in the GVCs of sophisticated products (e.g. electronics and automobile) which, as I will show later, have much higher economic potentials than low-skilled tasks in traditional sectors. However, on the other hand, the vertical upgrading turns out to be much less common. There is little relatedness between low-skilled and medium/high-skilled tasks. The relatedness remains rather low even for different skilled tasks within a same industry, this is especially the case in sophisticated manufacturing like electronics and automobile. While some case studies describe how an enterprise or a regional industry cluster upgrades towards higher value-adding activities along the value chain (see, e.g. Humphery and Schmitz 2002, Humphrey and Memedovic 2003), the skill upgrading seems to be rather difficult for a country as a whole. This result echoes the findings from Lemoine and Ünal-Kesenci (2004) and Jarreau and Poncet (2012) who show that the gains from assembly trade in China is largely limited within the processing exporter themselves, and their participation in GVCs does not contribute much to technology progress nor growth in the rest of the economy. Hence GVC participation does not guarantee a sustainable growth path in the long-run, and proper policy stimulations might still be necessary to achieve economy-wide vertical upgrading.

Another interesting finding is that various tasks in utility, logistics and transport sectors are closely related with manufacturing tasks. This highlights the complementarity between tasks in manufacturing and services, and suggests that a pro-trade business service sector can be developed in the early stage of development, which will continue to play a supportive role in the future progress of export upgrading. This important type of complementarity can never be revealed in a product space approach.

To provide further evidence that my task space approach is useful in structural change analysis, I also test whether the relatedness index is relevant for the actual changes in each country's export structure. Using probit regressions, I show that the probability in developing a comparative advantage in a new task is significantly positively related with its proximity with those tasks that in the initial period the country already has a comparative advantage in. This result is robust to alternative specifications, therefore confirms the validity of my task space in analyzing structural change as a dynamic process.

The rest of the paper is organized as follows: in section 2 I introduce the concept of relatedness in tasks and compare my relatedness with the current literature. Section 3 describes the methodology and the data I use in deriving value-added export of tasks. Then I explore the “task space, i.e. the structure of relatedness in section 4, followed by section 5 which illustrates actual changes paths in task space for countries at different level of development. Section 6 tests the power of the relatedness index in predicting actual directions of structural change. Section 7 concludes and discusses the policy implications from my research.

2 Co-occurance in RCA as a Measure for Task Relatedness

I build my task relatedness indices using the method by Hidalgo *et al.* (2007) which traces the relatedness of two products through the correlation in their revealed comparative advantage (RCA). In this paper, I will trace the relatedness in tasks instead of products.

The relatedness between two tasks x and y, denoted as $\phi_{x\to y}$, is defined as the probability that a country has a comparative advantage in task y, conditional on it having already a comparative advantage in task x:

$$\phi_{x\to y} = \text{Prob}(RCA_{i,y} > 1 | RCA_{i,x} > 1). \tag{1}$$

This conditional probability can be empirically calculated by:

$$\phi_{x\to y} = \frac{\text{Number of Countries with } RCA_y > 1 \text{ and } RCA_x > 1}{\text{Number of Countries with } RCA_x > 1}. \tag{2}$$

Similar as in Hidalgo *et al.* (2007) and Neffke *et al.* (2008, 2011), I calculate this ϕ statistic for each year in the period of 1995 to 2009 for which the data are available, and take the simple average of $\phi_{x\to y}$ across years to proxy the relatedness between each pair of tasks.

The idea behind this relatedness measure is intuitive. A high $\phi_{x\to y}$ means that a task y is frequently found to be a comparative advantage in the countries that also have a comparative advantage in x, which reveals that the socio-economic environment that nurses task x also fits the requirement for task y, therefore y will evolve and develop in the country with a higher probability. In the terminology of economic geography literature the co-occurrence of revealed comparative advantages in two tasks suggests that there are so-called “forces of agglomeration” that glue them together (Marshall

1920, Ellison *et al.* 2010). Two tasks may have similar requirements on workers, share common technology and intermediate inputs, have backward/forward linkages within a supply chain, and other sorts of synergy effects such that the development of one task also reinforces the growth of another. It can also be the case that the existence of a particular task is a pre-requisition for another. For example, ICT related services are unlikely to develop without the construction of a good network infrastructure.

The derivation of task relatedness requires the data on revealed comparative advantage indices. In this paper I construct the index in a similar way as the conventional RCA proposed by Balassa (1965), but I derive the RCAs using value-added export data on tasks instead of the gross exports of products. The RCA index for a task x in country i is defined as follows:

$$RCA_{i,x} = \frac{VAE_{i,x}/\sum_x VAE_{i,x}}{\sum_i VAE_{i,x}/\sum_{i,x} VAE_{i,x}}, \tag{3}$$

The interpretation of value-added based RCA index remains the same as the conventional RCA. The nominator measures the share of task x in country i's exported value-added, while the denominator captures the share of x in the value-added trade of the world. Therefore, a larger-than-unity $RCA_{i,x}$ means that country i exports a higher share of x relative to the world average, which implies that i has a revealed comparative advantage in task x.

As already discussed in the introduction, the value-added export based RCA is a clearer and more meaningful measure for the comparative advantages and henceforth task relatedness in globalised production. And compared with the relatedness in industries or products, the decomposition of value-added exports into the contributions by tasks at different skill levels provides more information on the actual activities that are taking place in an economy. To give an example, when one overlooks the skill levels and focuses on the export of products, textile and electronics may show up as highly related since textile industry and the final stage in electronics production (i.e. assembly) have actually quite similar requirements for labour, and indeed the largest exports of both electronics and textile show up in South East Asia and China. However, the relatedness between textile and electronics products in this case does not imply that textile is highly related with the "stereotype" high-tech activities in electronics (like the design and manufacturing of silicon microchips)[3]. It remains unknown whether the participa-

3. Neffke *et al.* (2008) shows another example in which the relatedness in industry becomes confusing. They have detected significant relatedness between "other textile" and "other food" industries with medicines. But their detailed exploration reveals that the factories producing bandages (classified as other textile) usually produce some basic disinfectants that can be combined with their bandages, while the production of spices and herbs are classified as other food but they may also have medical use. This

tion in low-skilled tasks of the electronics industry will induce the development in these higher skilled tasks.

It should be noticed that this measure of relatedness is based on the observed co-occurance in RCA which captures all potential forces of task relatedness. It is an empirical similarity measure and remains silent on the exact channel through which two tasks are related. There are also other researches aiming at specific pre-defined channels of relatedness between economic activities. Neffke and Henning (2013) use the Swedish official registers on employment to trace the movements of individual employees across different industries, which generates insights into the similarity of labour skills by each industry. Conley and Dupor (2003) use the U.S. input-output table and derive two measures of economic distance between industries: one measure investigates the similarity in two industries' input structure, and another focuses on whether two industries have backward/forward input-output linkages.

A final comment is needed concerning the asymmetry of task relatedness. Note that in most cases $\phi_{x\to y}$ does not equal $\phi_{y\to x}$, while Hidalgo *et al.* (2007) use a symmetric measure constructed from taking the minimum of relatedness from two directions, i.e. $\phi_{x,y} = \phi_{y,x} = \min(\phi_{x\to y}, \phi_{y\to x})$. I use the asymmetric measurement since the difficulty in developing a task x conditional on y is in general not the same as the development of y conditional on x. Two directions of relatedness should be treated separately as argued in Nedelkoska *et al.* (2015) who analyse the relatedness between occupations; it is relatively easy for a person to switch from a tough job to an easier one while the reverse is usually not true. The same argument holds in the process of structural change at country level. For instance, oil refinery and oil-related chemistry are quite likely to evolve in the countries that have many oil fields (i.e. a strong oil extraction industry). On the reversed direction, a majority of developed countries have refineries using imported oil, but it is impossible for them to develop a comparative advantage in oil extraction tasks because most of them do not have oil fields anyway. By taking the minimum of $\phi_{x\to y}$ and $\phi_{y\to x}$ as in Hidalgo *et al.*, the "difficult direction" of the transition counts, therefore the difficulty of upgrading from oil extraction to processing is largely overestimated. From an analytical point of view, it is more suitable to keep the asymmetric relatedness of both two directions, and use the relevant index according to research necessity. Policy makers are usually only interested in the path of upgrading from current activities to new and more preferred ones. In such case the difficulty

type of nuisance relatedness between industries can be severe when there is a high degree of international production fragmentation, making it more important to splitting the industrial production to tasks at different skill levels and investigate the tasks' relatedness.

associated with the "downgrading" path is not of particular interest[4].

3 Deriving the Value-added Export by Tasks

To calculate the relatedness between each pair of tasks, I need to derive first the value-added export (VAE) of tasks. Formally, a task is defined as the set of activities in an industry (or a service sector) that are performed by workers at a certain level of skill (low, medium- and high-skilled), according to the worker's educational attainment. Consider a task x referring to the activities in an industry a using type-f labour. Following Johnson and Noguera (2012) I define VAE of x as the value-added created by type-f labour in a that is ultimately absorbed as final demand in foreign countries.

The derivation of VAEs from a country i follows, briefly speaking, a backward-tracing strategy. The starting point is the final demand of all countries *other than* i itself. Then I use international input-output tables to decompose these foreign final demands into the value-added generated by ultimate factors from different country/industries that are required in the production. The part of value-added generated by tasks in country i is therefore the value-added exports by i that satisfies the definition above.

The formal steps of the VAE derivation are illustrated as follows. Assume there are m countries in the world and each country has n industries. The global input-output structure is described by a so-called technical matrix $\boldsymbol{A}$ of the size $mn \times mn$, with the

4. Hidalgo *et al.* (2007) are in favour of the symmetric relatedness indicators and in the online appendix they discuss about the potential flaws in using the asymmetric one. When a good z is exported only by one country, the asymmetric relatedness $\phi_{z \to z'}$ will, by construction, equal one for all z' that this country has a comparative advantage. Since one is the upper bound for the relatedness indices, it misleadingly implies a easy structural change from z to all those z'. They claim that this problem is mitigated by taking the minimum of $\phi_{z \to z'}$ and $\phi_{z' \to z}$. However, taking the minimum generates a misleading result from another direction. Namely, $\phi_{z \to z''}$ equals zero for all z'' that this country has a comparative disadvantage, and this zero relatedness "overrides" the non-zero $\phi_{z'' \to z}$ in taking the minimum, implying that the product z and all those z'' are completely unrelated. In general, this is a statistical small-sample problem which occurs only when the comparative advantage of a good z is observed in very limited number of countries. Under such circumstances, the relatedness $\phi_{z \to z'}$, i.e. the probability of having a comparative advantage in z' conditioning on z, cannot be accurately estimated using the equation (2) due to the lack of observations. Hidalgo *et al.* (2007) use detailed product classification at 4-digit level, so it is quite possible that some products are only exported by a single or few countries. This is, however, not likely to take place in this paper since I use 2-digit industrial classification. According to the data in 2009, a task appears at minimum in 6 countries as a comparative advantage, and more than 90% of the tasks appear as a comparative advantage in 10 or more countries.

following structure:

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1m} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & & \boldsymbol{A}_{2m} \\ \vdots & & \ddots & \vdots \\ \boldsymbol{A}_{m1} & \boldsymbol{A}_{m2} & \cdots & \boldsymbol{A}_{mm} \end{bmatrix}. \tag{4}$$

Each block $\boldsymbol{A}_{ij}$ is an $n \times n$ matrix, capturing the traded intermediate inputs from country i that are used by the production in j; a typical element $A_{(i,a),(j,b)}$ stands for the value of intermediate goods produced by country i's industry a that are used in producing \$1 output by the industry b in j. The matrix $\boldsymbol{A}_{ii}$ on the diagonal stands for the domestic input-output structure within each country i.

I use $y_{i,a}$ to denote the value of gross output from country i's industry a. And country j's demand for the final good produced by industry a of country i is denoted by $d^j_{i,a}$. In matrix form, the total production from the world and the final demand of each country are represented in mn-element column vectors $\boldsymbol{y}$ and $\boldsymbol{d}^i$ as follows:

$$\boldsymbol{y} = \begin{bmatrix} y_{1,1} \\ y_{1,2} \\ \vdots \\ y_{1,n} \\ y_{2,1} \\ \vdots \\ y_{2,n} \\ \vdots \\ y_{n,m} \end{bmatrix}, \qquad \boldsymbol{d}^1 = \begin{bmatrix} d^1_{1,1} \\ d^1_{1,2} \\ \vdots \\ d^1_{1,n} \\ d^1_{2,1} \\ \vdots \\ d^1_{2,n} \\ \vdots \\ d^1_{n,m} \end{bmatrix}, \quad \boldsymbol{d}^2 = \begin{bmatrix} d^2_{1,1} \\ d^2_{1,2} \\ \vdots \\ d^2_{1,n} \\ d^2_{2,1} \\ \vdots \\ d^2_{2,n} \\ \vdots \\ d^2_{n,m} \end{bmatrix}, \quad \cdots, \quad \boldsymbol{d}^n = \begin{bmatrix} d^n_{1,1} \\ d^n_{1,2} \\ \vdots \\ d^n_{1,n} \\ d^n_{2,1} \\ \vdots \\ d^n_{2,n} \\ \vdots \\ d^n_{n,m} \end{bmatrix}. \tag{5}$$

The final demand of the whole world is by definition the summation of final demands from all individual countries, i.e. $d^W_{i,a} \equiv \sum_j d^j_{i,a}$ and in matrix form $\boldsymbol{d}^W \equiv \sum_j \boldsymbol{d}^j$.

Equality holds that $y_{i,a} = d^W_{i,a} + \sum_{j,b} y_{j,b} A_{(i,a),(j,b)}$, i.e. the value of total gross output from country i's sector a equals the total final demand for its products, plus those outputs used as intermediate inputs by all countries and industries. This input-output structure of the whole world can be written in a compact matrix form as:

$$\boldsymbol{y} = \boldsymbol{d}^W + \boldsymbol{A}\boldsymbol{y}, \tag{6}$$

where $\boldsymbol{A}\boldsymbol{y}$ captures the total usage of intermediate inputs. Re-arranging this equation using the Leontief Inverse (Leontief 1949) gives the relationship between total production and total final demand:

$$\boldsymbol{y} = (\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{d}^W. \tag{7}$$

To investigate the value-added export from country i, the total demand $\boldsymbol{d}^W$ can be decomposed into two components: final demand of country i itself, $\boldsymbol{d}^i$, and final demand of all foreign countries $\boldsymbol{d}^{-i} = \sum_{j \neq i} \boldsymbol{d}^j = \boldsymbol{d}^W - \boldsymbol{d}^i$. Correspondingly, the total gross production of the world can also be split into two components:

$$\boldsymbol{y} = (\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{d}^i + (\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{d}^{-i} = \boldsymbol{y}^i + \boldsymbol{y}^{-i} \tag{8}$$

the part $\boldsymbol{y}^i$ is a vector representing the total gross outputs in each country/industry that are required to satisfy the final demand of country i, and $\boldsymbol{y}^{-i}$ is the production requirements for satisfying the world final demand outside of country i. Each element in the latter term $\boldsymbol{y}^{-i}$ that belongs to the industries of country i, for example $y^{-i}_{i,a}$, captures the gross output of industry a of country i that is required to satisfy the final demand from abroad.

We are interested in the value-added export contributed by tasks at each skill level f (low-, medium- and high-skilled). I use $\eta_{i,a,f}$ to denote the value added by type-f labour in producing \$1 gross output in sector a of country i. The value-added export by each activity $x \equiv \{a, f\}$ in country i is therefore given by:

$$VAE_{i,x} \equiv VAE_{i,\{a,f\}} = \eta_{i,a,f} y^{-i}_{i,a}. \tag{9}$$

Subsequently, the RCA of each task and the task relatedness indices can be calculated by equation (3) and then (2).

To empirically compute VAE by each task, this paper uses the recently developed World Input Output Database (WIOD, Timmer *et al.* 2015) as the primary data source. The WIOD dataset provides the time series of multi-regional world input-output tables (WIOTs) which is available from 1995 to 2011. WIOTs tell about the production structure of the world, namely the level of production of each industry and the domestic and imported intermediate input usages for all countries/industries. It also provides information on final consumption of domestic or imported goods/services by each country. The database consists of 40 countries that cover more than 85% of world output; it includes most of the developed countries, but also a majority of emerging economies like Eastern European countries and the "BRIC" (Brazil, Russia, India and China), which are the most interesting for structural upgrading researches. The estimates of the economic structure related with the rest-of-world (RoW) countries are also provided by the WIOTs, such that the production and consumption for the whole world is covered.

WIOTs are constructed from official data sources from multiple agents, including the national supply and use tables provided by national statistical bureaus and UN trade data at disaggregated levels. And the dataset covers the trade in services and intangibles based on multiple sources like Eurostat and WTO. The WIOTs are constructed at 2-digit industry level (ISIC Rev.3), and data across countries are harmonized such that most of the industries in different countries are comparable. Note that tasks in several industries are excluded from this analysis, either because the industry has large discrepancies in registration standards across countries, or because the industry serves mainly domestic final use only and associate with neither direct nor indirect export (the RCA index cannot be calculated for non-exporting tasks). The excluded industries and their ISIC codes are: Hotel and restaurants (H), Other transport and operation of tourist agents (63), Public administration and defense (L), Education (M), Healthcare (N), Personal and community services (O), and Self-employed persons (P)[5]. In total I include 28 out of 35 ISIC industries.

To decompose sectoral value-added output into the contribution by activities at different skill levels, I further use the data on the skill structure of employment in each country/industry (or service), from the Socio-economic Accounts (SEA) in the WIOD project. In particular, I use the shares of factor income earned by low-, medium- and high-skilled labour in each country/industry, according to educational attainment[6]. This supplement dataset is available from 1995 to 2009 and I focus my research on this period. The time-series nature of WIOD also provides a possibility for me to trace the actual dynamics in industrial structure in the 15-year time period for these 40 countries, through which I can systemically investigate the power of relatedness indices in predicting the actual directions of export structure change.

As a comparison with gross export data used in the conventional RCA index, VAE properly controls for the position of a task in the global value chain (see also Johnson and Noguera 2012 and Koopman *et al.* 2014). A country located in down-stream of the GVC has a higher level of gross export but it is caused by a higher intermediate to gross output ratio. The conventional RCA based on gross export does not adjust for this, and will exaggerate the comparative advantage of countries in the down stream. In contrast, when calculating the VAEs the higher usage of intermediate inputs translates by construction into a lower ratio of $\eta_{i,a,f}$, and all values related with intermediate inputs are effectively netted out. Furthermore, conventional RCA tends to systemically

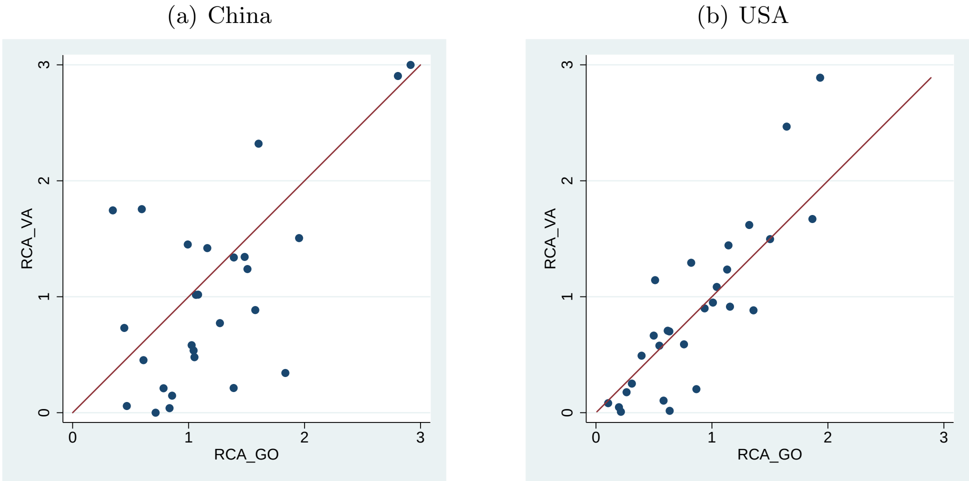
5. The value-added created by tasks in these industries are excluded from analysis, but the values of intermediates used by them are included. For example, if the Japanese healthcare sector uses the medicines produced in the U.S., these values still count and will appear as the VAE by the U.S. chemistry sector.

6. The detailed classification of skill level is according to the International Standard Classification of Education (ISCED). Low-skilled workers refer to employees with lower-secondary school, or lower level of education. Medium-skilled workers are those who have high school education. High-skilled workers have college or higher level of education.

understate the comparative advantages of some sectors in the countries where they rely mostly on indirect export, since indirect exports simply do not show up in gross export data. Instead, VAE includes both directly exported value-added as well as the value-added that are further used by other domestic exporting industries; the algorithm in calculating VAE identifies where the value-added is created and consumed regardless of how many intermediate processing stages being taken.

To see the differences between gross export based and VAE based RCAs, I show the value-added RCA at industry level because gross export based RCAs cannot trace the comparative advantages down to the task level. Figures 1a and 1b show the discrepancy between RCA_{GO} and RCA_{VA} for each industry in China and the US in the year 2009; the differences between these two measures are quite large.

Figure 1: Comparison of Gross Export and Value-added Export based RCA Index



Notes: Each dot represents two measures of RCA at ISIC rev. 2 industry level for the year 2009. The horizontal axis plots the value-added based RCA of an industry, and the vertical axis plots the conventional RCA based on gross export. The straight line is $y = x$; dots that are further away from the line imply a larger discrepancy between two RCA measures. Based on WIOD database and author's own calculation.

Table 1 provides more detailed information about different measures of RCA for several manufacturing industries in China and the US in 2009. The most notable differences are observed in electronics and chemistry industries, which provide typical examples for two sorts of distortions discussed above. For the chemistry industry in China, the VAE based RCA is almost twice of the conventional one. It is due to the fact that 87% of output from Chinese chemistry industry is used further as intermediate inputs by other Chinese firms domestically, which is much larger compared with 54% in the U.S. The majority exported value created by the Chinese chemistry industry therefore goes indirectly through other exported products and it explains the discrepancy between two

RCA measures.

Table 1: Different Measures of Revealed Comparative Advantages

I - China

Industry	RCA_{GO}	RCA_{VA}	Ratio	II/GO (%)	DUI (%)	RCA_{L}	RCA_M	RCA_H
Textile	3.00	2.91	1.03	79.5	56.1	5.05	2.56	0.41
Leather	2.90	2.80	1.04	80.2	47.1	4.48	2.20	0.33
Chemistry	0.54	1.04	0.51	79.4	87.0	1.64	1.04	0.33
Plastic	1.34	1.48	0.91	81.3	81.9	2.78	0.91	0.18
Metal	0.77	1.27	0.61	80.3	89.9	1.94	0.77	0.25
Machinery	1.02	1.06	0.96	77.0	50.3	1.91	0.80	0.27
Electronics	2.32	1.60	1.45	83.9	44.1	3.05	1.37	0.45
Automobile	0.45	0.61	0.74	80.5	51.9	1.28	0.52	0.18

II - United States

Industry	RCA_{GO}	RCA_{VA}	Ratio	II/GO (%)	DUI (%)	RCA_{L}	RCA_M	RCA_H
Textile	0.25	0.31	0.81	56.3	53.7	0.16	0.50	0.72
Leather	0.08	0.11	0.78	51.7	54.5	0.05	0.19	0.30
Chemistry	0.95	1.01	0.94	66.5	53.8	0.09	0.67	1.59
Plastic	0.70	0.63	1.11	67.5	76.3	0.21	0.78	0.70
Metal	0.59	0.76	0.78	66.4	87.5	0.29	1.04	0.82
Machinery	1.08	1.04	1.04	58.0	27.5	0.14	0.75	0.86
Electronics	0.88	1.36	0.65	43.5	32.7	0.22	1.21	2.68
Automobile	1.23	1.13	1.09	72.5	34.5	0.18	0.84	1.36

Notes: The column "Ratio" is $\mathrm{RCA}_{GO}/\mathrm{RCA}_{VA}$, "II/GO" is the cost share of intermediate inputs in gross output, "DUI" is the share of gross output that are used domestically by other industries as intermediate inputs. Source: Author's own calculation based on WIOD dataset.

Electronics gives a different picture. The gross export based RCA shows that China has a strong comparative advantage with a RCA of 2.3, while U.S. has a comparative disadvantage. In contrast, in the VAE based RCA Chinese electronics industry scores only a moderate 1.6, and in the U.S. electronics turns into a comparative advantage. The discrepancy between gross export and VAE based RCAs exceeds 50% (0.88 versus 1.36). The divergence between two measures is explained by the intermediate input to gross output ratio. Chinese electronics industry is located in the down-stream part of the value chain; it uses a relatively high level of intermediate inputs that comprises of 84% of the gross output from the industry. The high intermediate to output ratio inflates gross export statistics and leads to an over-optimistic estimate of RCA. On the other hand, the U.S. electronics industry adds more than 55% values on its own in 2009, and has the lowest intermediate to output ratio among all 40 countries covered by this research.

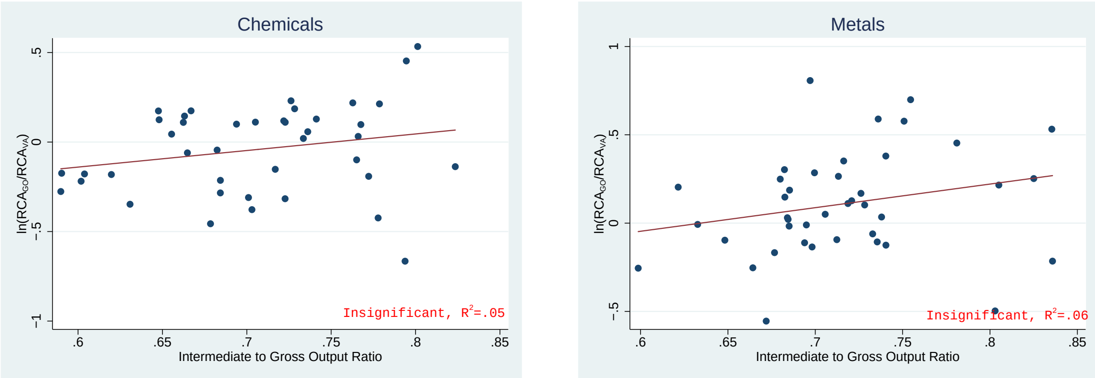
To give a more systemic illustration of the bias of RCA caused by share of intermediate in gross output, in figure 2 I plot the bias of RCA (measured by $\ln(RCA_{GO}/RCA_{VA})$) against the intermediate to gross output ratio for a selected number of industries across countries in the year of 2009. A positive correlation is found for all these sectors, namely the conventional RCA systemically overstates the role of downstream countries. The correlation is positive but insignificant for the industries that produces mainly intermediate goods (chemical and metal industries), while the upper-sloping trend is highly significant in industries producing sophisticated goods (electronics and automobile) where the intermediate to gross output ratio itself explains more than 50% variation in the RCA bias.

My result on RCA confirms the findings by Koopman *et al.* (2012, 2014) that conventional RCA tends to be systemically overestimated (underestimated) for countries/industries located in down-stream (upper-stream) of the value chain. However, it is important to notice that the VAE measure used in this paper is different from Koopman *et al.*'s domestic value-added in export (DVA). The measurement of DVA is easier as it requires only the domestic input output table of each country (Los *et al.* 2016). It measures the domestically added values embedded in each country's gross export, but lacks the ability in identifying the exact tasks that generate the value-added in export. Industry level VAEs, on the other hand, can be decomposed into the contributions by tasks at different skill levels. Therefore, the comparative advantage can be measured at task level which is not possible when using gross export or DVA data. This is shown in the last three columns in Table 1. Task level VAE gives more precise information on the actual activities taken place in each country. China and the U.S. have quite similar value-added based RCAs for electronics and chemicals at industry level, and China performs even a bit better than the US. Given the fact that electronics and chemicals goods are generally considered as advanced products with high technology intensity, it seems to suggest that China and the U.S. are at a similar level of development. But when we decompose the VAEs into contributions by different skilled tasks, it becomes clear that China's comparative advantages in these "advanced products" are mainly within the low-skilled segment, while the U.S.'s comparative advantages are in the high-skilled tasks.

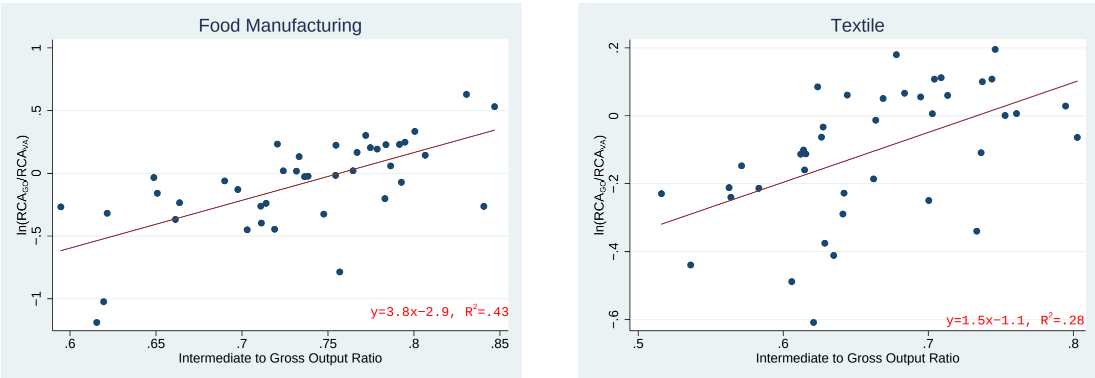
Another important advantage of the VAE measure is that it identifies the indirected exported value-added from (non-traded) business support services which are becoming increasingly important in the economy; this is not possible when using DVA. While the previous literature about relatedness and structural change focuses on the industrial sector, I am also able to derive the relatedness of business services tasks with the tasks in the manufacturing sector, to investigate the role of services in structural change. Indeed

Figure 2: The Bias of Gross Output Based RCAs and the Link with Intermediates to Gross Output Ratio

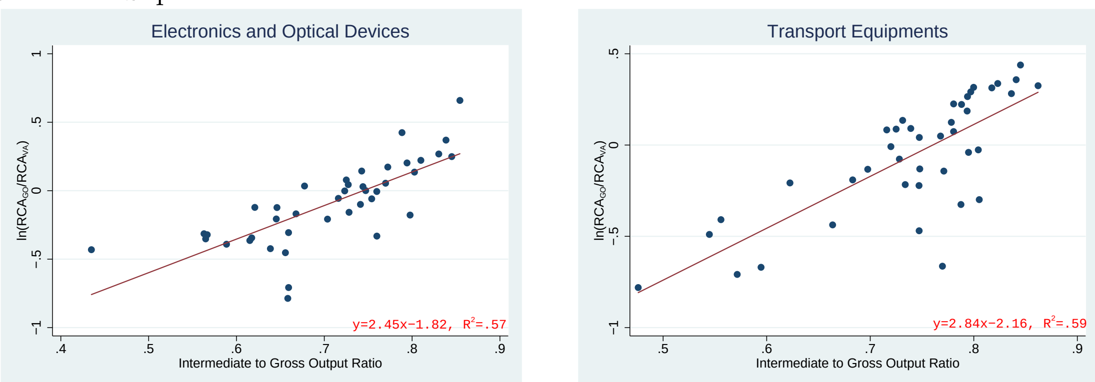
Sectors Related with the Production of Materials



Sectors of Simple Products



Sectors of Sophisticated Products



Notes: author's own calculation based on data for 40 countries in WIOD database for the year of 2009. Each dot represent a country. The bias in RCA is represented by the ratio between gross and value-added export based RCAs (in logarithm). Red line is the linear fit curve for the scatter plot based on simple OLS regression. The OLS coefficients are reported for the industries where the slope coefficient is significant at 0.01 level. R^2 corresponds to the fitness of OLS.

as I will show in the next two sections, tasks in utility and logistics sectors play a quite important role in the upgrading paths.

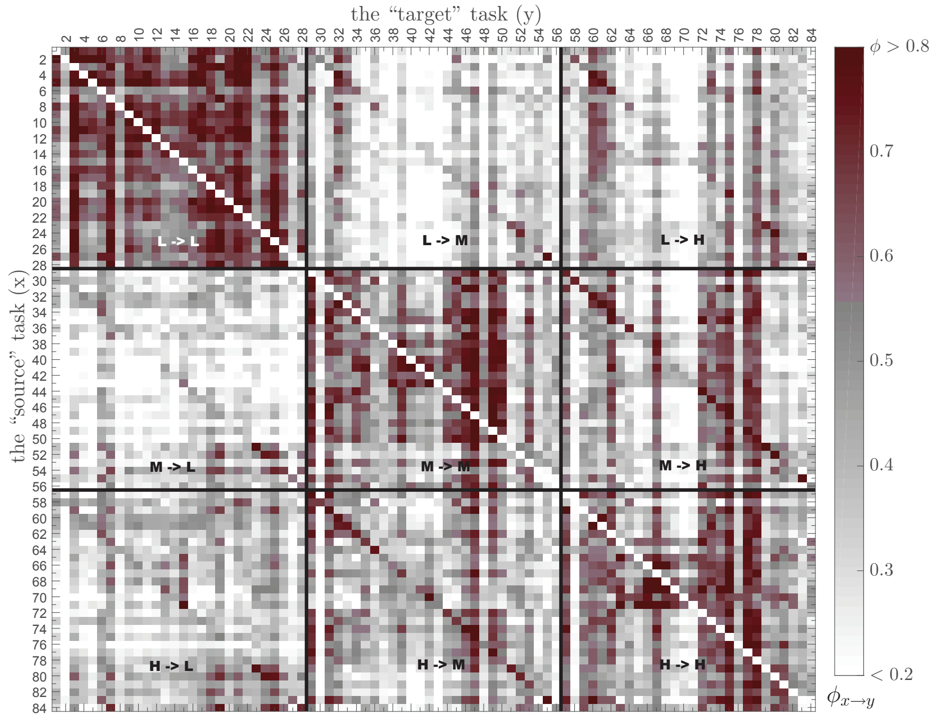
4 The Structure of the Relatedness between Tasks

Using the value-added export of tasks, I derive my "task space" which is an 84×84 matrix containing the relatedness between each pair of tasks. The structure of relatedness will be explored in this section. To have a quick view on the relatedness between tasks, the task space is presented as a heatmap in figure 3. I number tasks from 1 to 84, and organize them into three groups according to their skill levels. Tasks numbered 1 to 28, 29 to 56, and 57 to 84 correspond to low-, medium- and high-skilled ones; detailed correspondence table can be found in the appendix. The heatmap contains 84×84 small squares, the colour of a square on row x column y represents the value of $\phi_{x \to y}$. A darker colour stands for a higher value of relatedness index; note that in the graph a task associated with a dark-coloured column has strong relatedness with the other tasks, and therefore this task will be relatively easy to start with. To highlight the upgrading in skill, the heatmap is divided by black lines into nine big blocks according to the tasks' skill levels. And to further facilitate the visualization, the colour of a small square is dyed with red if the associated $\phi_{x \to y}$ is above 0.55 which is the criteria for closely related tasks as set in Hidalgo *et al.* (in my research about 22% relatedness indices fit this criteria).

It is evident from the heatmap that the three big blocks on the diagonal a have high density of deep coloured squares, indicating that a task is, in general, more closely related with other tasks at the same skill level. This is especially the case among low-skilled tasks as shown by the deepest color in the block $L \to L$ on the north-west corner. The three big blocks above the diagonal are the most interesting for developing countries, since the relatedness represented by squares in these blocks are associated with the likelihood of the structural change towards high-skilled tasks. The upgrading from low-skilled to higher skilled tasks should be difficult according to my task space, as the $L \to M$ and $L \to H$ blocks contain only a few dark columns. The upgrading from medium-skilled to high-skilled tasks is relatively easier, especially in many business services sectors (72 to 80), however the relatedness between medium- and high-skilled tasks in many modern and sophisticated manufacturing sectors (65 to 71) are still low.

An important target of the paper is to evaluate the difficulty in horizontal and vertical types of upgrading. The horizontal upgrading, i.e. the shift of low-skilled

Figure 3: Heatmap of Bilateral Relatedness Index beteen Tasks



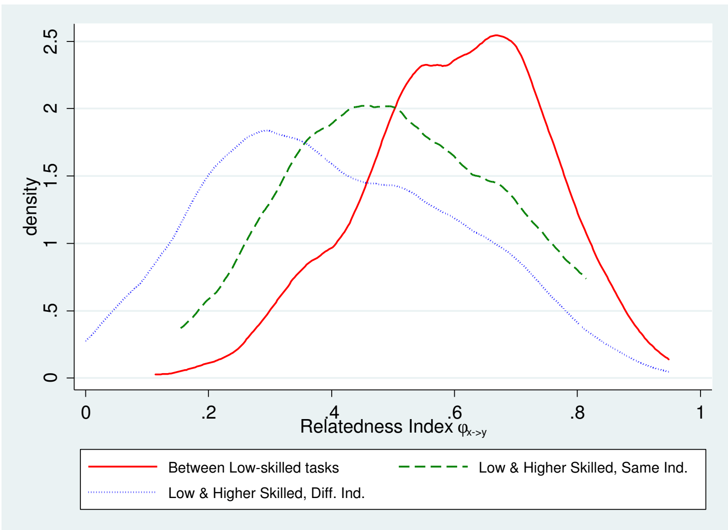
Notes: Based on the task space, i.e. 84×84 matrix of $\phi_{x\to y}$. Darker color is associated with a higher relatedness value. Blocks correspond to $\phi_{x\to y} > 0.55$ are indicated by red.

employment towards the sophisticated value-chains, will be relatively easy according to the high relatedness between all kinds of low-skilled activities. For example, the relatedness $\phi_{x\to y}$ representing the re-allocation from low-skilled textile task to electronics task is 0.67, and to low-skilled automobile task 0.63. It means for countries that already have a comparative advantage in low-skilled task of textile, two-third of them have a comparative advantage in low-skilled tasks in electronics (or automobile) as well, which shows that low-skilled tasks in sophisticated products' value chains have a good fitness with the socio-economic conditions of countries at a relative early stage of development. Therefore, for underdeveloped countries that have not yet integrated in the GVCs, the horizontal reallocation of low-skilled workers from traditional sectors towards GVCs will be a smooth process with a high probability of success.

What about the difficulty in vertically upgrading for less developed countries, namely the upgrading along the value chain towards higher skilled tasks within an industry? The relatedness structure suggests that vertical upgrading is on average more difficult than the horizontal upgrading, but the difficulty differs a lot across industries.

To investigate within-industry skill upgrading, one should look at the squares on the diagonal lines in the $L \to M$ and $L \to H$ blocks, as these squares represent the values of $\phi_{x \to y}$ such that the task x and y belong to a same industry (i.e. $y - x$ equal to 28 or 56, see also the task code table in the appendix). In the heatmap the diagonal lines are darker than other squares in these two blocks, indicating that the task upgrading towards higher-skilled activities within an industry is relatively easier than the upgrading towards other higher-skilled activities in different industries. But it is still not very apparent to compare the vertical upgrading's difficulty with horizontal upgrading in the heatmap. For a clearer comparison, in figure 4 I draw the density plots for three sets of relatedness indices that are associated with different types of upgradings. The solid line represents the density of relatedness between two low-skilled tasks (i.e. horizontal upgrading), while the dashed line for the relatedness between low-skilled tasks and the medium- and high-skilled counterparts within each industry (i.e. vertical upgrading). As shown in this figure, the relatedness for horizontal upgradings are on average higher than the relatedness for vertical upgrading; in the former case 64% relatedness indices are larger than 0.55, while 40% for the later. The dotted line is the distribution of $\phi_{x \to y}$ regarding cross-industry skill upgrading, which is distributed much further to the left.

Figure 4: Distribution of Relatedness Index for Different Types of Upgrading



Notes: Kernel density distributions of relatedness $\phi_{x \to y}$, for three kinds of upgradings from low-skilled tasks.

The difficulty in vertical upgrading is also predicted to be quite heterogeneous. Vertical upgrading seems to be relatively easy in the traditional manufacturing sector (textile, leather and shoes, wood products), the corresponding relatedness between low-skilled tasks and the medium- and high-skilled counterpart is around is around 0.7 (among the top 10% percentile in all relatedness indices). However, in more "attractive" manufacturing sectors, namely in the GVCs of sophisticated products, vertical upgrading turns out to be much more difficult. The $\phi_{x \to y}$ regarding the upgrading from low- to medium-skilled tasks in machinery, electronics and automobile sectors are only 0.38, 0.33, and 0.32 respectively, which are lower than the medium of all relatedness indices. It suggests that horizontal and vertical upgradings in the sophisticated GVCs are very different processes. The participation in low-skilled tasks in the GVCs is relatively easy due the strong relatedness with other low-skilled tasks in traditional sectors, but the participation does not seem to guarantee the vertical upgrading to happen spontaneously.

An interesting finding is that the task space reveals the complementarity between tasks in manufacturing and several business service sectors. I find the high-skilled tasks in several business service sectors show a high relatedness with their low-skilled counterpart, with a $\phi_{x \to y}$ larger than 0.55. These sectors are electricity, gas and water supply (73), inland transport (78), water transport (79), sale & maintenance of vehicles (75), renting of machinery & equipment and other business services (83), and even air transport (80). Note that all these services are related with utility and logistics that support domestic and international trade. More importantly, the medium- and high-skilled services tasks are also highly related with many other low-skilled tasks in the manufacturing sector, as indicated by the dark colored columns in the $L \to M$ and $L \to H$ blocks. This type of relatedness is unlikely to be an outcome of labour reallocation, but rather the complementarity between manufacturing and trade. This pattern is more clearly illustrated in the network analysis in the next section, in which I will also show that the development in higher-skilled trade supporting tasks is empirically observed in developing countries actual upgrading paths (for example in China).

5 Paths of Structural Change in the Task Space

What are the directions for structural upgrading suggested by the task space? In this section I am going to transform the relatedness indices into a network graph that depicts possible structural upgrading paths. And then I will also visualise the actual observed export structural change in the network graphs for countries at different levels of development, to investigate whether my task space is in line with the structural upgradings in the real world.

5.1 Structural Upgrading and the Economic Potential of Tasks

Structural upgrading is a directed progress in which a country moves towards more attractive tasks, i.e. the tasks that have higher potentials for economic growth. To analyze the direction of structural upgrading and evaluating the benefits from a certain upgrading path, it is important to first measure the economic potentials of tasks.

Higher-skilled tasks are usually viewed to better than the low-skilled one within the same industry. However, the economic potentials cannot be easily compared in such an adhoc way for same skilled tasks across industries, and it can be the case that certain lower-skilled tasks have better economic potentials than some higher-skilled tasks in other industries. These call for an objective measure for each task's economic potential. In this paper, I follow Hausmann, Rodrik and Hwang (2007) and use a modified PRODY index to measure the economic potentials of tasks. The PRODY index of a task x is defined as follows in this paper:

$$PRODY_x = \frac{\sum_i w_i y_i RCA_{i,x}}{\sum_i w_i RCA_{i,x}}, \tag{10}$$

where y_i is the per-capita income of country i, and w_i is an importance weight for each observation and here I use the country's economic size (national GDP)[7].

In table 2 I list five tasks that have the highest/lowest PRODY over the period 1995 to 2009. PRODY is a revealed measure of economic potential and can be interpreted as the income level of a representative country that export this task. To put it in another word, a high PRODY means that the task is usually the comparative advantage of developed countries with a high level of income per capita.

PRODY provides an ordering for the attractiveness of tasks in structural change. The magnitude of PRODY is also meaningful, and the benefit from a particular path of structural change can be proxied by the difference between the PRODY indices of two

7. Income per capita the and national GDP are the real (per-capita) output-side GDP, fetched from the Penn World Table 8.1 (Feenstra, Inklaar and Timmer 2015). The PRODY index used here is a weighted PRODY which differs from the original one $\frac{\sum_i y_i RCA_{i,x}}{\sum_i RCA_{i,x}}$ as in Hausmann, Rodrik and Hwang (2007). The original PRODY index can be viewed a weighted average of income per-capita across countries, and the weight for each observation is the comparative advantage of each country. If a task frequently has a high RCA in richer countries, the per-capita incomes of those richer countries get a higher weight and it translates into a high PRODY index for this task. However, small countries usually have a much higher degree of specialisation and have more extremely high values of RCA indices than larger countries. Effectively small countries are over-represented, making the original PRODY index less meaningful. For example, the top 3 most attractive tasks according to the original PRODY are low-skilled finance, low-skilled business services, and medium-skilled finance. I use economy size (national GDP) as the importance weight w_i for each country, and the weighted PRODY gives a more meaningful ranking of tasks' economic potentials.

Table 2: Tasks with the highest and lowest PRODY index

Top 5		Bottom 5	
Task	PRODY	Task	PRODY
High-skilled Electronics	32064	Low-skilled Agriculture	8357
Med-skilled Air Transport	31039	Low-skilled Mining	8990
High-skilled Publishing	30784	Low-skilled Leather & Shoes	9775
High-skilled Business Services	30661	Low-skilled Textile	10093
High-skilled Automobile	30073	Low-skilled Oil Refinery	11527

Source: Author's own calculation based on WIOD. Averaged PRODY across 1995 to 2009.

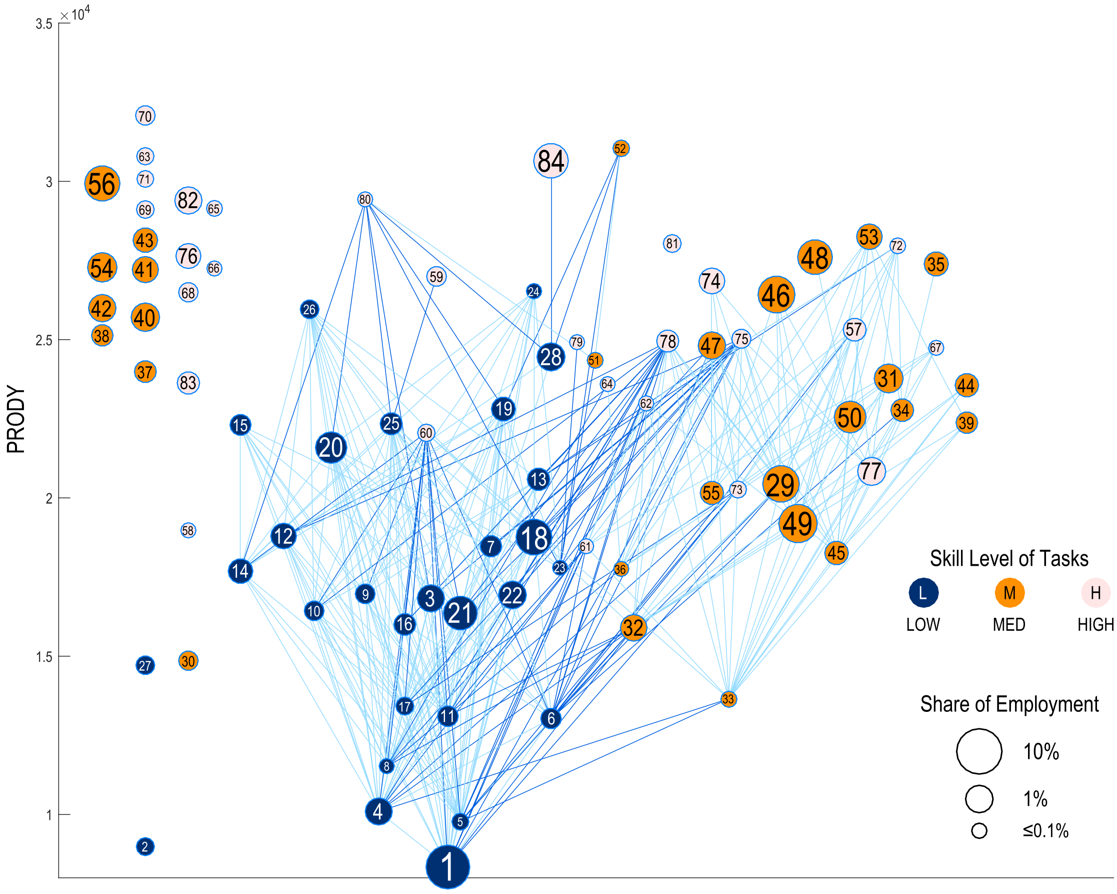
tasks. To give an example, the PRODY indices for low- and medium-skilled tasks are 10093 and 15898 in the textile sector, and are 17693 and 25977 in electronics. So if we do not consider the difficulty in structural change directions for a moment, starting from the low-skilled tasks in textile, the vertical upgrading is less attractive than the horizontal upgrading towards electronics. If after the integration with the GVCs the country can further achieve the vertical upgrading in the electronics industry as suggested by Lin (2012), a much larger gain is expected for the horizontal upgrading towards electronics.

5.2 Possible Upgrading Paths in a Network Graph

With a proper measure for the economic potentials of tasks, the structural upgrading process of a country can be intuitively described in a so-called "jumping monkey" analogy as discussed in Hidalgo *et al.* (2005). Each task can be viewed as a tree, and its economic potential is represented by the number of fruits that the tree is able to produce. The employment in a task is then in analogy to the monkeys living in a particular tree, and the progress of structural upgrading can be viewed as the monkeys jumping towards the trees that provide them with the most number of fruits. But monkeys must take the distance between trees into consideration before making a jump, as a monkey can only jump towards the trees that are close enough to the tree it currently lives in (i.e. two tasks must have sufficiently high relatedness).

The possible paths for structural change are therefore in analogy to a "map" that plots the possible jumping routes between the trees, and I illustrate it by the mean of a network graph as in figure 5. A network consists of *nodes* and *edges*. In the network of my task space, nodes represent the tasks at various skill levels. Different skilled tasks are represented by different colors, and their size indicate the average share of employment of tasks in the economy (for the 40 WIOD countries in 1995, unweighted). The edges represent the feasibility of the directions of structural change. The structural change from one task to another is said to be possible if there is an edge linking these two nodes. In this paper I use the criteria by Hidalgo *et al.* (2007); two nodes are linked by an edge if the associated $\phi_{x \to y}$ is greater than 0.55.

Figure 5: Network Graph of the Relatedness in the Task Space



The nodes are vertically positioned in the graph according to their PRODY index, such that in figure 5 the nodes from the bottom to the top are in an ascending order according to their economic potentials[8]. To reduce the complexity of the graph and to focus more on the process of structural upgrading, I only draw the edges that are associated with a upgrading direction where the "target" task's PRODY index is at least 20% higher than the "source" (i.e. $\text{PRODY}_y/\text{PRODY}_x \geq 1.2$). I also exclude the

8. The horizontal axis of my network graph does not have a particular meaning; the horizontal positions of nodes are chosen to reduce the number of crossings of edges and to make the graph more readable. Note that my network graph is different from the product space network as in Hidalgo *et al.* (2007); they do not visualize the economic potentials of products and both horizontal and vertical positions of nodes are determined by an algorithm to minimize the number of crossings of edges.

relatedness with the tasks in which less than 10 countries have a comparative advantage to increase the reliability of the visualization (see the discussion in footnote 4, in total 3 nodes are affected). An illustration on how the network graph is created can be found in the appendix. Note that there are also many "isolated tasks" that do not have any edge with others; I position these nodes on the left side of the graph.

Figure 5 confirms the findings based on the heatmap that horizontal upgrading is much more likely than the vertical one. Nodes of low-skilled tasks are heavily linked with each other, and at the same time low-skilled tasks in different industries show a large heterogeneity in terms of their potentials for economic growth. The low-skilled tasks in sophisticated products (e.g. electronics and automobile) have a PRODY index around $20,000, which is almost twice of the PRODY of traditional low-skilled activities like in agriculture and textile (around $10,000). This shows that horizontal upgrading is a feasible option and has an important growth enhancing effect for for countries at an early stage of development.

On the other hand, There is only a limited number of edges that span from low-skilled tasks towards higher skilled ones (highlighted by a darker color in figure 5), and mostly towards the medium- and high-skilled tasks in utility and logistics sectors (47, 51, 52, 73, 75, 79, 80, etc) but not many towards higher-skilled manufacturing. The opportunity for vertical upgrading is rather low; only upgrading towards higher skilled tasks in traditional manufacturing seems to be promising (in textile and leather products, 32, 33, 60, 61). The medium- and high-skilled tasks in modern manufacturing sectors (37 to 44 and 65 to 72) turn out to have the highest level of PRODY index, therefore the vertical upgrading in these industry is expected to yield a much higher gain for an economy in the future. However there is little relatedness between these tasks with other tasks in the rest of the economy. For example, in electronics industry the relatedness between low- and medium tasks (14 and 42) is only 0.33, and 0.45 between medium- and high-skilled (70). When looking at the relatedness index with other tasks, only the high-skilled task in automobile (71) has a relatedness higher than 0.55 with the high-skilled electronics task. However they have very similar levels of PRODY which are among the top of all tasks. Therefore this route is not very relevant with structural upgrading, and is not an option for developing countries that do not have any comparative advantage in high-skilled sophisticated manufacturing.

Similar as in the heatmap, the network graph also shows that medium- and high-skilled tasks in utility and logistics-related business services have a higher probability to evolve even during the early stage of development. There might be two possible reasons behind the relatedness between these services tasks and low-skilled tasks in manufacturing industries. Firstly, utility and logistics sector complement the low-skilled manufacturing tasks by reducing energy and transportation costs. Compared with many medium- and high-skilled manufacturing tasks which are frequently embedded in light-weighted but valuable products, low-skilled manufacturing tasks are usually associated with physical processing and are low in their value-added. The costs spent on energy and logistics constitute therefore a relatively larger share in the effective price. More efficient utility and logistics therefore reduce the effective costs in delivering the low-skilled tasks to the world market, which increase the competitiveness of low-skilled manufacturing firms. And on the other hand, an abundant supply of low-skilled labour means the skills of personnels in logistics and utility sectors can be utilized at a larger scale. This increases the ability and willingness for firms in these services sectors to pay a high wage and keep a handful high-skilled staffs for the key positions. It in turn attracts highly educated young people to enter these sectors.

The second reason is the relatively low propensity of offshoring for tasks in utility and trade supporting sectors. In principle the high-skilled tasks in these sectors are remotely transmittable (for instance, the scheduling in logistics), but the nature of these sectors makes the solutions provided outside of the country less competitive. Different from ICT or financial services where remote work from another country will not significantly dampen the quality of output being delivered, many of the high-skilled tasks in utility and logistics require location and relationship specific knowledge that are frequently tacit (for example, knowing the regional traffic or power grid conditions). This might create a natural barrier against foreign competitors, such that the high-skilled tasks have a higher chance to develop even when a country is still at a relatively low level of development.

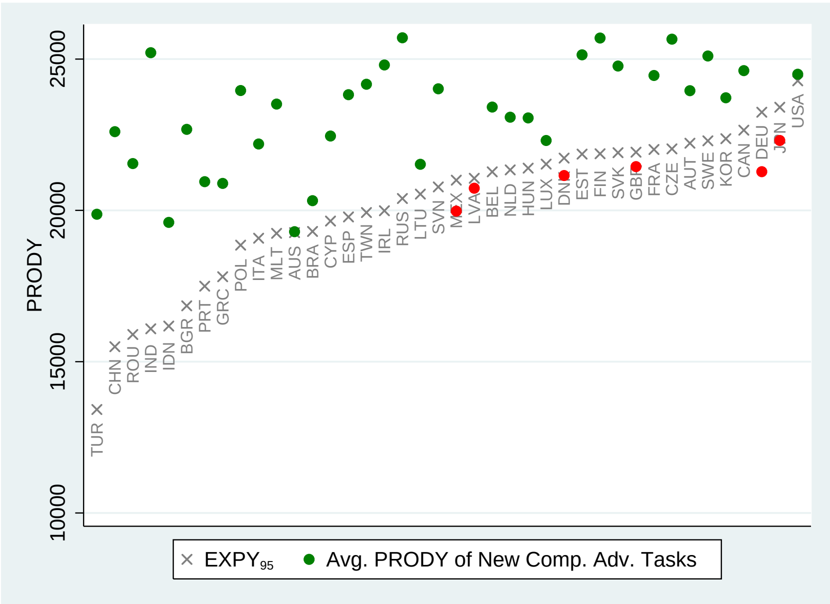
5.3 Actual Structural Upgrading Paths in the Task Space

To see whether the actual directions of export structural change are aligned with my task space, in this subsection I will visualise the changes in comparative advantages in the task space network, for countries at different level of development during 1995 to 2009. In this period, most countries have experienced an upgrading in their economic structure. This is shown by comparing the PRODY index of newly developed comparative advantages with the country-wide economic potential in the initial period. The economic potential of a country can be proxied by the so-called EXPY index (Hausmann, Hwang and Rodrik 2006), which is a weighted average of PRODY and the weights are the shares of value-added export of each task for each country i:

$$EXPY_i = (\sum_x VAE_{i,x} PRODY_x)/\sum_x VAE_{i,x}. \tag{11}$$

In figure 6 I show the initial EXPY for each country in 1995, and the average PRODY of the tasks that the country has gained a comparative advantage from 1995 to 2009. It shows that the new comparative advantages have a higher PRODY in 34 out of 40 economies, and in the six exceptions the PRODY of new comparative advantages is lower than the national EXPY but the difference is small.

Figure 6: Average PRODY of New Comparative Advantageous Tasks in Each Country



Notes: Dots indicate the average PRODY index of tasks that a country newly developed a comparative advantage during the period of 1995 to 2009 (simple average). Crosses indicate the level of EXPY of each country in 1995. EXPY captures the economic potential of a country; countries gaining new comparative advantages in tasks with a PRODY higher than its initial EXPY therefore has experienced economic upgrading. Both PRODY and EXPY are based on the data in 1995.

While most countries are experiencing a structural upgrading, the newly developed comparative advantages differ across different types of countries. Figures 7a to 7c depict the changes in task-level RCAs from 1995 to 2009 for three subsets of countries in the WIOD database: Non-European developing countries, Eastern European countries, and developed countries[9]. In each figure, plots I and II show the comparative advantages of each country group in 1995 and 2009, plot III identifies the new and disappeared com-

9. Non-European developing countries: Brazil, China, India, Indonesia, Mexico, Russia and Turkey. Eastern European Countries: Bulgaria, Cyprus, Czech, Estonia, Greece, Hungary, Lithuania, Latvia, Malta, Poland, Romania, Slovakia, and Slovenia. Developed countries: other countries in the WIOD database.

parative advantages of the region, and plot IV identifies the tasks with largest increase and decrease in the value of their RCA.

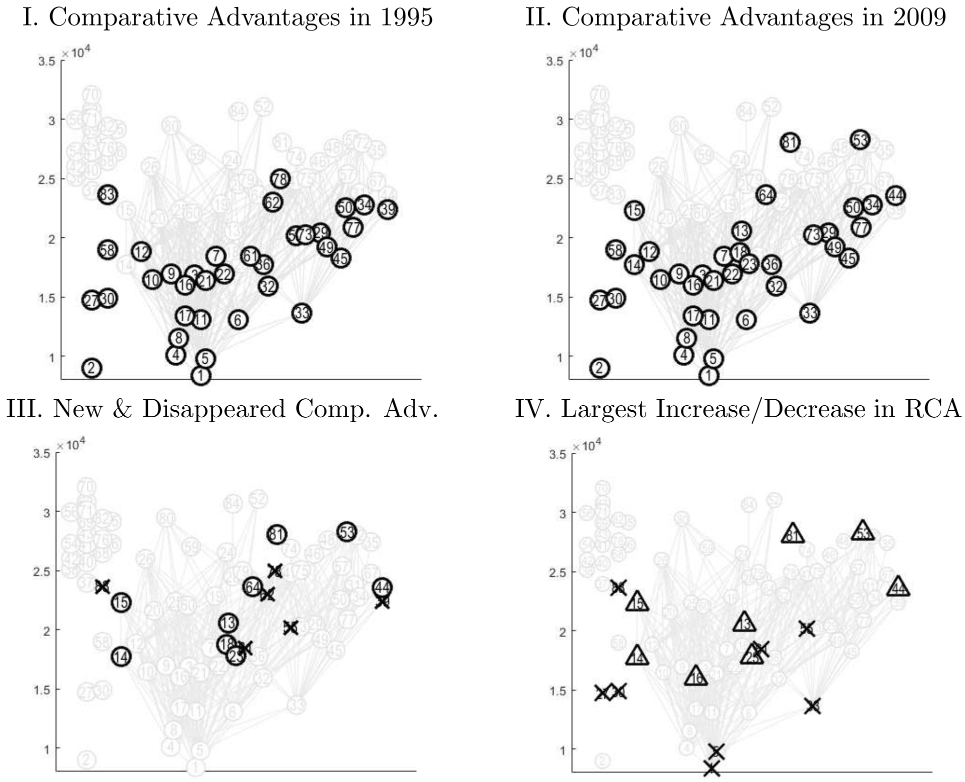
Non-European developing countries have on average the lowest level of per-capita income among all countries covered by the WIOD dataset. Their comparative advantages are mainly in low-skilled tasks in both 1995 and 2009. The tasks that have the quickest development are those low-skilled tasks with relatively high economic potentials, most notably in the Machinery, Electronics and Automobile industries (code 13 to 15). New comparative advantages also arise in some non-isolated medium- and high-skilled tasks. However, there is no notable development in the isolated medium- and high-skilled tasks although they have the highest growth potential according to the PRODY index. Low-skilled agricultural task remains as a comparative advantage in 2009, but the decrease in its RCA value is one of the largest among all tasks. These patterns in combination suggest that structural change paths of non-European developing countries can be described as horizontal upgrading, i.e. the shift of low-skilled employment from traditional activities towards low-skilled tasks in the value chains of sophisticated products.

Eastern European countries are at a higher level of development. In 1995 they already have many comparative advantages in several non-isolated medium- and high-skilled tasks that are on the upper-right corner of the plot. During 1995 to 2009, a considerable share of comparative advantages in low-skilled tasks has disappeared, and the Eastern European countries were developing their new advantages in medium- and high-skilled tasks including some of the isolated ones. The advanced countries already have the comparative advantages in high- and medium-skilled tasks in 1995, mostly among the favourable but isolated tasks on the upper-left corner. During the 15 years developed countries keep consolidating these comparative advantages; they did not have the comparative advantages in low-skilled tasks from the beginning but the largest decreases in RCA are still observed from those low-skilled ones.

China is a fast emerging economy and a large destination of offshoring in the recent decades, and it has achieved quick integration with the world market. I illustrate the export structural change of China separately in figure 7d using the same set of plots as 8a-c. Similar as other non-European developing countries, the comparative advantages of China mainly lie in low-skilled tasks. Plot 8d-III shows that China has gained quite many new comparative advantages at medium-skilled level including those desirable but isolated ones. It suggests that China has overcome the lack of relatedness and has undergone vertical skill upgrading at a quite early stage of development, considering its low income level in 1995 (real income per capita \$3,445 at 2005 constant price according to PWT 8.1, the second lowest among all countries in the WIOD database and comparable

Figure 7: Dynamics of Comparative Advantages from 1995 to 2009

a. Non-Europe Developing Countries



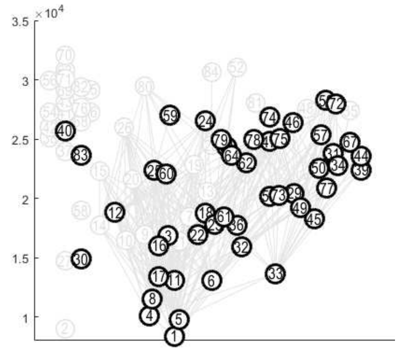
Notes: Figure 7a-c are based on the average of RCAs in of each country group, weighted by the country sizes in 1995 (or 2009). Figure 7d is based on the RCAs in China. In each figure, the rounds and crosses in plot III indicate new and disappeared comparative advantages. And in Plot IV, triangles and crosses show the tasks with top 10% increases/decreases in their RCA values.

to countries like Egypt and Sri Lanka). But when we look at the largest changes in the RCA values, plot IV shows China's upgrading pattern looks much like other developing countries, namely the RCAs rise the most in low-skilled tasks in sophisticated products (13 to 15), and the trade supporting services (23 and 51 on water transport, 73 on utility, and 81 on telecommunication), and the largest decreases in RCA are found in agriculture and traditional manufacturing.

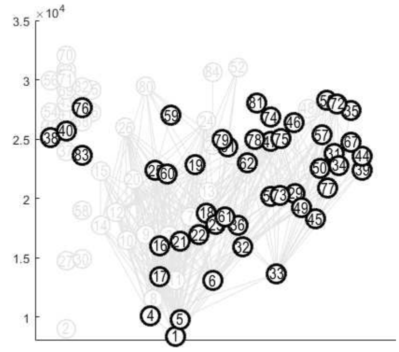
To see whether China has undergone a rapid vertical upgrading along the value chain, it is worthwhile to have a closer investigation on the changes in the RCA values of each task from 1995 to 2009, as shown in table 3. The low-skilled tasks in electronics has undoubtedly the largest increase in the RCA, and medium-skilled electronics task has also a notable growth in RCA of more than 0.6 and it turned into a comparative advantage by 2009. For other two isolated tasks that China has gained comparative advantage, i.e. medium-skilled tasks in Chemistry and Finance (37 and 54), their increases

7b. Eastern European Countries

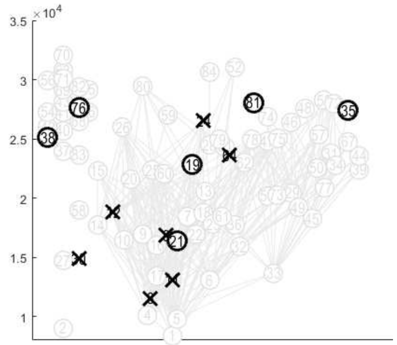
I. Comparative Advantages in 1995



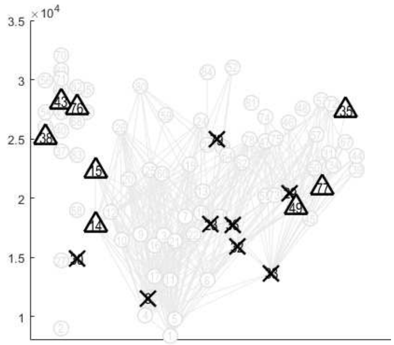
II. Comparative Advantages in 2009



III. New & Disappeared Comp. Adv.

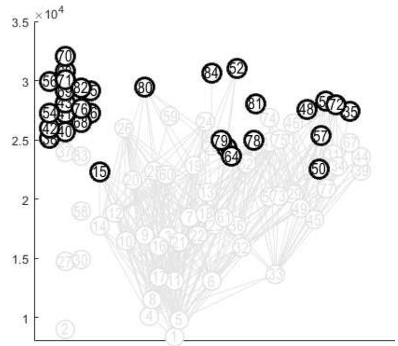


IV. Largest Increase/Decrease in RCA

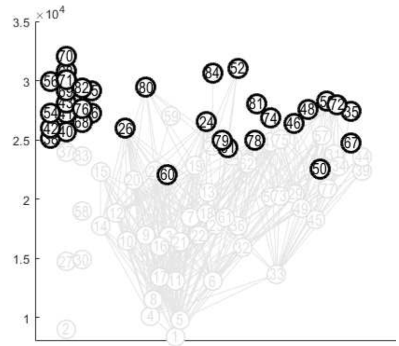


7c. Developed Countries

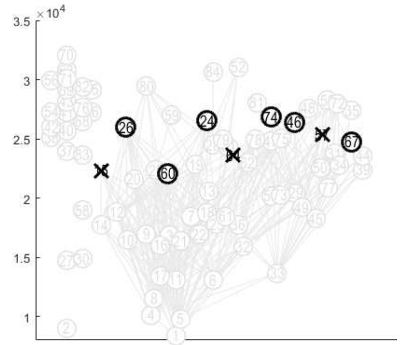
I. Comparative Advantages in 1995



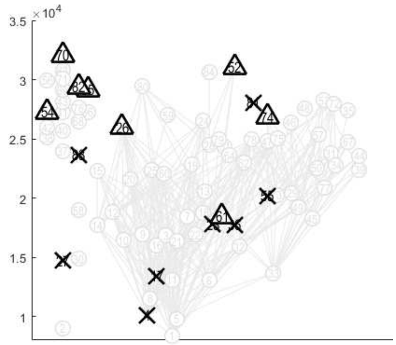
II. Comparative Advantages in 2009



III. New & Disappeared Comp. Adv.

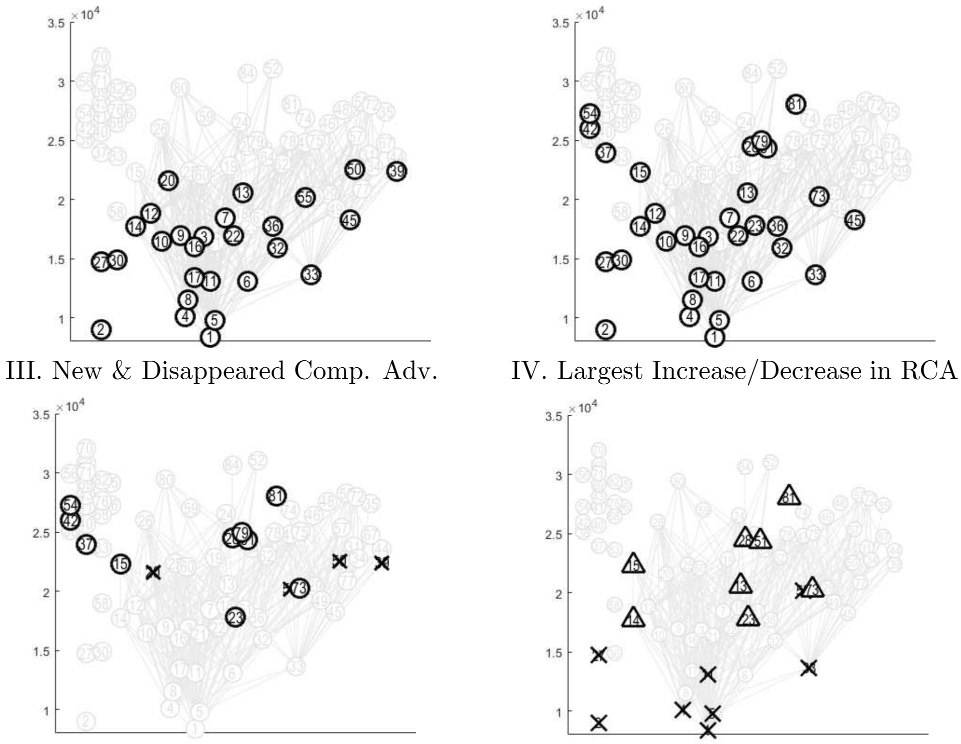


IV. Largest Increase/Decrease in RCA



7d. China

I. Comparative Advantages in 1995 — II. Comparative Advantages in 2009 — III. New & Disappeared Comp. Adv. — IV. Largest Increase/Decrease in RCA



in RCA are much more moderate. As a comparison, many tasks in utility and trade supporting sectors not only switched into the status of a comparative advantage, but also have quite large increases in their RCAs. There are also other service tasks that have a large rise in the RCA during 15 years but did not reach comparative advantage yet, for example the high-skilled tasks in retailing and air transport. And when we investigate the skill structure of employment in each Chinese industry [10], the share of medium- and high-skilled labour in Chinese electronics industry has increased by 6.1 percentage points during the 15 years, which is faster than, but not very different from other manufacturing sectors. It seems that the new comparative advantage in medium-skilled electronics is an outcome of a rapid expansion of the total size of the Chinese electronics industry rather than within-industry skill upgrading. On the other hand, the largest within-industry skill upgrading is observed in retailing and wholesaling sectors (about 20 percentage points increase in the share of medium- and high-skilled labour in total employment).

10. The data are provided in the socio-economic account of WIOD. See Table 7 in the appendix for details.

Table 3: Revealed Comparative Advantages of Tasks in China, 1995 and 2009

Sector \ Skillness	1995 L	1995 M	1995 H	2009 L	2009 M	2009 H
Agriculture	6.25	0.65	0.05	4.85	0.88	0.02
Mining	2.66	1.55	0.15	1.63	1.09	0.33
Food Manufacturing	1.35	0.70	0.10	1.69	0.89	0.24
Textile	5.76	2.47	0.20	5.05	2.56	0.41
Leather Products	5.71	2.75	0.19	4.48	2.20	0.30
Furniture	2.71	0.78	0.08	2.99	0.88	0.17
Paper and Publishing	1.50	0.50	0.04	1.70	0.58	0.09
Refinery	1.72	1.54	0.28	1.51	1.48	0.67
Chemistry	1.25	0.77	0.13	1.64	1.04	0.33
Rubber and Plastic	2.34	0.80	0.09	2.78	0.91	0.18
Non-metal Minerals	4.05	1.34	0.13	3.08	0.93	0.16
Metal	1.92	0.79	0.14	1.94	0.77	0.25
Machinary	1.02	0.41	0.08	1.91	0.80	0.27
Electronics	1.67	0.74	0.15	3.05	1.37	0.45
Automobile	0.40	0.18	0.04	1.28	0.52	0.18
Other Manufacturing	1.24	0.37	0.04	1.84	0.54	0.10
Utility	2.51	1.33	0.34	2.57	2.00	1.25
Construction	0.84	0.35	0.06	0.69	0.28	0.10
Sale of Automobile	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Wholesaling	1.04	0.75	0.29	0.62	0.79	0.80
Retailing	0.31	0.35	0.08	0.22	0.37	0.19
Inland Transport	1.73	1.10	0.18	1.25	0.70	0.24
Water Transport	0.95	0.44	0.17	2.58	1.48	1.01
Air Transport	0.26	0.51	0.44	0.23	0.50	0.94
Post and Communication	0.16	0.28	0.29	0.29	0.76	1.25
Finance	0.14	0.98	0.16	0.23	1.22	0.42
Real Estate	4.60	1.80	0.25	3.12	0.99	0.22
Business Services	0.51	0.16	0.04	1.42	0.42	0.23

Source: Author's own calculation based on WIOD dataset. L, M, H stands for the low-, medium- and high-skilled tasks in each industry. Note that China doesn't have a separated industry classification for "Sale and Repair of Motor Vehicles and Motorcycles; Retail Sale of Fuel" (ISIC code 50).

6 Testing the Role of Relatedness in the Actual Directions of Structural Change

To provide further evidences that countries upgrade through the task space towards highly related tasks, I am going to show that a task's proximity with current comparative advantageous tasks has significant predicting power on whether the country will develop a new comparative advantage in this task.

The concept of economic activity relatedness has been used in various researches, however there is still a lack of systemic econometric tests in the literature. Recently Kali *et al.* (2013) measures the proximity between the comparative advantageous products of a country with the set of products that it does not export yet, and they use it to reflect the country's distance with future comparative advantages. They show a country has a higher probability in experiencing a period of growth acceleration if its initial economic structure has closer link with comparative disadvantageous products.

My research has a different focus, and I am interested in a systemic test on the paths of structural change, to see whether relatedness play a significant role in determine the new comparative advantages in the future. Formally, I am going to test the following hypothesis:

> *The probability that a country gains comparative advantage in a task is positively associated with the proximity of this task with the current export structure of the country.*

I use the following indicator to measure the proximity of a task x with the current export structure in each country i:

$$\widetilde{\phi}_{i,x} = \sum_{y \in \mathrm{Adv}_i} s_{i,y} \phi_{y \to x}, \qquad \mathrm{Adv}_i\text{: the set of tasks in country } i \text{ with } \mathrm{RCA}_y > 1. \tag{12}$$

In the equation, $s_{i,y}$ is the share of employment by task y in country i (measured by the number of working hours which is provided in the socio-economic account of WIOD); all variables are based on the information in the initial period. This measure reflects an "overall" proximity of a task x with the initial export structure in the country, taking both the strength of relatedness (i.e. $\phi_{y \to x}$) and the relative size of related comparative advantageous tasks ($s_{i,y}$) into consideration. It can be viewed as a "support" for a new task from the economic structure of country i: a task gets a stronger support if it has higher relatedness with the comparative advantageous tasks that have larger shares of employment in the economy.

This proximity indicator is different from Kali *et al.* and Hidalgo *et al.*, who separately measures the strength of relatedness and the size of economy related with the new task. Their strength measure is the maximum relatedness of a task with current comparative advantageous tasks, i.e. $\phi_{i,x}^{\max} = \max_{y \in \mathrm{Adv}_i} \{\phi_{y \to x}\}$, which I will include in the robustness check. Another measure they use is constructed as $\omega_{x,i} = \sum_{y \in \mathrm{Adv}_i} \phi_{y \to x} / \sum_y \phi_{y \to x}$ which they refer to as the "density" of relatedness. The target of this variable is to measure the "percentage of neigbhouring space around the new task that is already developed in a country" (Hidalgo *et al.* 2005), however it effectively assign a same weight to each product (or task) code and does not take the actual structure of economic activities into consideration[11].

I use the data from 1995 to 2009 and I focus on the tasks that have an $\text{RCA}_{95} < 0.75$ in the initial year. The tasks whose initial RCAs are close one are excluded in the benchmark analysis, since whether the task is a comparative advantage or disadvantage (i.e. whether RCA is larger or smaller than 1) can be sensitive to the fluctuations in trade and measurement errors in these cases. The growth of RCA from 0.75 to above 1 is large enough to ensure it is not merely due to these distortions[12].

To provide a quick diagnostics on the role of proximity in determine the probability of gaining a new comparative advantage, in figure 8 I plot the share of tasks that has gained a comparative advantage for the observations with different levels of $\widetilde{\phi}_{i,x}$. Tasks with the highest level of proximity has almost 10 times probability in becoming a new comparative advantage compared with those lowest.

Formally, I test the hypothesis using the following probit regression:

$$\begin{aligned} &\text{Prob}(\text{RCA}_{i,x,09} > 1|\text{RCA}_{i,x,95} < 0.75) \\ &\quad = \Phi(\text{Const.} + \beta_1\widetilde{\phi}_{i,x,95} + +\beta_2 Potential_{i,x,95} + \beta_3\text{RCA}_{i,x,95} + \varepsilon_{i,x}), \end{aligned} \tag{13}$$

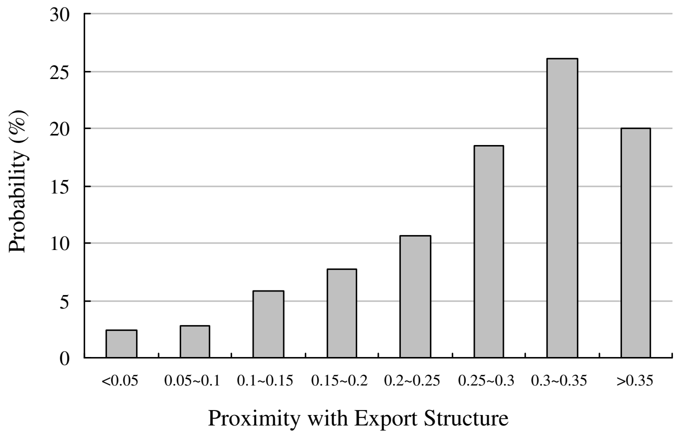
where $\Phi(.)$ is the cumulative distribution function of the standard normal distribution. The core parameter in the regression equation is β_1. A positive significant β_1 implies that the probability in gaining a comparative advantage is positively associated with the proximity of the task with the comparative advantageous tasks, therefore relatedness has predicting power on actual direction of structural change.

I include two important control variables. Since structural change is in general an upgrading process towards the activities with a higher growth potential than the economic average level, I control for the task's attractiveness as a target for structural change by its economic potential relative to the country's average level. The variable $Potential_{i,x,95}$ is constructed as the logarithm of $PRODY_x/EXPY_i$; it is positive (neg-

11. To see that, consider four products. The relatedness between products 1 & 3 and between 2 & 4 is 0.9; the relatedness between any other pairs is 0.1. Assume country A and B have comparative advantage in product 1 and 2. These two products make up, say, 60% and 30% of total employment in A, and 30% and 60% in B. Intuitively, product 3 is more likely to develop in country A (and product 4 in B) since it is related with a larger share of the economy. The density measure as in Hidalgo *et al.* and Kali *et al.* does not use the information of employment structure and will give the relatedness density $\omega_3 = \omega_4 = (0.9 + 0.1)/(0.9 + 0.1 + 0.1) = 10/11$ for both product 3 and 4 in both country A and B.

12. Hidalgo *et al.* (2007) use 4-digit industry classification and focus on the observations with an initial RCA smaller than 0.5. I classify tasks at 2-digit level; at such an aggregated level it is more difficult for RCA to grow quickly, therefore a higher threshold 0.75 is chosen. Focusing on only RCA< 0.5 would result in a smaller sample size and more importantly, much fewer number of tasks had gain a comparative advantage in this subset which dampens the power of probit regression I use.

Figure 8: Probability of Gaining a Comparative Advantage (95–09)



Notes: Only observations with an initial $\mathrm{RCA}_{95} < 0.75$ are included. The vertical axis is the measured probability in gaining a comparative advantage, calculated as the number of observations that have gained a comparative advantage in the period 1995 to 2009, divided by the total number of observations in a certain range of $\widetilde{\phi}_{i,x}$ as represented on the horizontal axis. The variable $\widetilde{\phi}_{i,x}$ is the task's proximity with the initial export structure of a country, see equation 12.

ative) when the PRODY of the task is higher (lower) than the EXPY of the country. I also include the task's initial RCA value as a control variable, because if other things stay the same, it would be easier for a task to obtain the status as a comparative advantage (i.e. RCA>1) if it has a higher RCA at the beginning.

The regression results are reported in table 4; the "MEF" columns besides a probit regression refer to the marginal effects evaluated at the mean of independent variables. Column 1 is the baseline model corresponding to the regression equation 13 above. As expected, the coefficient β_1 is positive and strongly statistical significant, indicating that countries indeed tend to gain new comparative advantages in the tasks that has higher proximity with current comparative advantageous tasks. The effect is also economically important. The marginal effect for specification (1) shows that when the proximity with current comparative advantages increase by one standard deviation (around 8 percentage points), the probability in gaining a comparative advantage is predicted to increase by 1.7 percentage points. This effect is large in the view that only 7.7% tasks has gained a comparative advantage across the whole sample.

In specifications (1a) and (1b), I use $\phi_{i,x}^{\max}$ as an alternative measure for task's proximity with current export structure. When only $\phi_{i,x}^{\max}$ is included, the coefficient is positive significant, but the R^2 of the regression has a notable compared with (1). And when both proximity measures are included in (1b), only $\widetilde{\phi}_{i,x}$ is significant and the result is virtually the same as (1), indicating that there is no need to control for the maxi-

Table 4: Regression Results

I - Main Results

	Baseline		Use $\phi_x^{\max}$	Both	Symetric Relatedness		Developing Econ	
	(1)	MEF	(1a)	(1b)	(2)	MEF	(3)	MEF
$\widetilde{\phi}_{x,i}$	3.593	0.215		3.948	4.788	0.283	2.767	0.234
	0.625***	0.044***		0.843***	0.807***	0.057***	0.753***	0.067***
$\phi_{x,i}^{\max}$			1.363	-0.353				
			0.418***	0.560				
Potential_{95}	1.443	0.086	1.625	1.388	1.440	0.085	1.182	0.100
	0.223***	0.013***	0.234***	0.238***	0.222***	0.013***	0.275***	0.022***
RCA_{95}	3.300	0.198	3.128	3.342	3.289	0.195	3.205	0.271
	0.309***	0.026***	0.307***	0.317***	0.310***	0.026***	0.382***	0.039***
Const.	-3.649		-3.789	-3.506	-3.767		-3.360	
	0.218***		0.307***	0.312***	0.230***		0.261***	
R^2	0.260		0.236	0.260	0.262		0.248	
Obs.	1646		1646	1646	1646		843	
# Positive	131	(7.69%)	131	131	131		77	(9.12%)

II - Testing Vertical Upgrading

	M&H Tasks, All Observations				M&H Tasks, Developing Countries			
	(4a)	(4b)	(4c)	(4d)	(5a)	(5b)	(5c)	(5d)
$\widetilde{\phi}_{x,i}$	4.025	4.246	4.102		2.732	3.222	3.135	
	0.795***	0.815***	0.806***		0.988**	1.040**	1.036**	
Potential_{95}	1.499	1.615	1.553	1.380	0.989	1.175	1.155	0.973
	0.304***	0.318***	0.318***	0.308***	0.410*	0.432**	0.043**	0.422*
RCA_{95}	3.159	3.242	3.191	3.149	2.902	2.986	2.939	3.001
	0.347***	0.353***	0.351***	0.343***	0.433***	0.437***	0.434***	0.432***
RCA_L		-0.069				-0.103		
		0.052				0.068		
IsRCA_L			-0.076	0.021			-0.232	-0.077
			0.126	0.122			0.177	0.167
Const.	-3.680	-3.701	-3.688	-2.968	-3.170	-3.236	-3.217	-2.675
	0.263***	0.263***	0.263***	0.203***	0.327***	0.333***	0.331***	0.263***
R^2	0.208	0.210	0.208	0.171	0.173	0.180	0.178	0.154
Obs.	1098				541			
# Positive	107	(9.7%)			61	(11.3%)		

Notes: Significance levels: *: 0.1, **: 0.01, ***: 0.001. All specifications are probit regressions, the dependent variable is a dummy indicating whether the country has gained a comparative advantage in the task in the year of 2009. I use the observations with initial $\text{RCA}_{95} < 0.75$. Specifications (4) and (5) focus only on medium- and high-skilled tasks, and specifications (3) and (5) uses only the observations from developing countries (real GDP per capita in 2000 smaller than \$20,000). MEF is the associated marginal effect evaluated at mean, using the delta method. "# Positive" indicates the number of observations with a dependent variable equals one in the probit regression, and the term in the bracket indicates the percentage in the total number of observations. The row of R^2 is the pseudo-R^2 of the probit regressions.

mum strength of relatedness when $\widetilde{\phi}_{i,x}$ is in place. In column (2) I use the symmetric relatedness (i.e. $\phi_{x,y} = \min(\phi_{x\to y}, \phi_{y\to x})$ as in Hidalgo *et al.* 2007) to construct the proximity measure, and in in specification (3) I focus only on developing countries with real GDP per capita lower than $20,000 in 2000. In both cases I obtain a similar result as the benchmark. Noteworthy, there is no statistical difference in the marginal effects of $\widetilde{\phi}_{i,x}$ in (1) and (3), which indicates that task's proximity with the current export structure has a similar effect on the probability in gaining a comparative advantage for both developing and developed countries.

In the lower panel of table 4, I test the propensity of vertical upgrading. To be more specific, I test the hypothesis that whether the existence of a comparative advantage in a particular low-skilled task is positively correlated with the probability that the country gain a comparative advantage in medium- and high-skilled task within the same industry. Column (4a) is a replication of the baseline regression, but focus only on the new comparative advantages in medium- and high-skilled tasks. In (4b) I include the RCA of low-skilled task in the same industry as a control variable, and alternatively in (4c) I include a dummy indicating whether the low-skilled task is a comparative advantage of the country in 1995. However, the coefficients on low-skilled task's RCA are insignificant in both specifications, while coefficients of all other variables are almost the same as the (4a). It is possible that the effect of vertical relatedness is already controlled by including the "overall" proximity $\widetilde{\phi}_{i,x}$ in the regression, but even $\widetilde{\phi}_{i,x}$ is excluded, column (4d) shows that the existence of a comparative advantage in the low-skilled task of an industry still has no significant correlation with the probability in gaining comparative advantage in higher-skilled tasks. In column (5a-d) I replicate the probit regressions of (4a-d) but focus on the subsample of developing countries, the result still does not give any support for vertical upgrading.

To check the robustness that proximity with current export structure predicts future comparative advantages, in table 5 I perform regressions using several alternative specifications. Column (6) includes all observations that has an initial $\text{RCA}_{95} < 1$, and column (7) investigates the probability that the RCA value of a task has grown more than 0.25 between 1995 and 2009. The values of coefficients are not comparable with the baseline specification due to the different sample coverage and regression criteria, but in all cases of $\widetilde{\phi}_{i,x}$ is significant and the marginal effects are large in magnitude[13]. In column (8) and (9) I perform OLS regression on the changes in RCA between 1995 and 2009. The dependent variable in (8) is the difference $\text{RCA}_{09} - \text{RCA}_{95}$, and in (9) the relative growth of RCA $\ln(\text{RCA}_{09}/\text{RCA}_{95})$. There are few observations with extreme

13. For instance, one standard deviation increase in $\widetilde{\phi}_{i,x}$ is associated with $0.821 * 8 = 7.16$ percentage points higher probability that the RCA of a underdeveloped task increases by at least 0.25.

Table 5: Regression Results for Alternative Specifications

	All $RCA_{95} < 1$		RCA Growth> 0.25		Changes in RCA	
	(6)	MEF	(7)	MEF	(8)	(9)
$\widetilde{\phi}_{x,i}$	3.523	0.352	3.102	0.821	0.617	1.498
	0.514^{***}	0.056^{***}	0.451^{***}	0.119^{***}	0.085^{***}	0.202^{***}
Potential_{95}	1.610	0.161	1.336	0.354	0.266	0.842
	0.180^{***}	0.017^{***}	0.131^{***}	0.032^{***}	0.017^{***}	0.041^{***}
RCA_{95}	3.060	0.306	0.996	0.264	0.009	-0.667
	0.190^{***}	0.025^{***}	0.117^{***}	0.046^{***}	0.031	0.076^{***}
Const.	-3.568		-1.724		0.004	0.148
	0.171^{***}		0.109^{***}		0.017	0.042^{***}
R^2	0.307		0.125		0.167	0.261
Obs.	2024		1646		1630	1583
# Positive	271	(13.4%)	368	(22.4%)	-	-

Notes: Significance levels: *: 0.1, **: 0.01, ***: 0.001. Specification (6) and (7) are probit regressions; MEF is the associated marginal effect evaluated at mean, using the delta method, and "# Positive" indicates the number of observations with a dependent variable equals one in the probit regression. Column (6) is the replication of (1) but include all observations with $RCA_{95} < 1$. The dependent variable in specification (7) is a dummy indicating whether the RCA of the task has increased by more than 0.25. Specification (8) and (9) are OLS regressions using $\text{RCA}_{09} - \text{RCA}_{95}$ and $\ln(\text{RCA}_{09}) - \ln(\text{RCA}_{95})$ as dependent variables, respective. The observations in (8) and (9) are trimmed to be within the 0.5 to 99.5 percentile in their dependent variable to control for outliers. Column (6) and (7) report pseudo-R^2; (8) and (9) report adjusted R^2. Note that the number of observation differs in (8) and (9), this is because the dependent variable in (9) is not well defined if a task does not exist in a country in either 1995 or 2009 (i.e. RCA=0).

changes in their RCA indices, so I exclude these outliers and only include the observations whose change in RCA is within 0.5 to 99.5 percentile. Both specifications show that a task's proximity with export structure positively correlates with the growth in RCA index, which reconciles my hypothesis.

Lastly, there is are potential validity and endogeneity concerns on the core explanatory variable, i.e. the proximity measure $\widetilde{\phi}_{i,x}$. Note that $\widetilde{\phi}_{i,x}$ is constructed using the relatedness index which is ultimately based on the pool of RCA values from 1995 to 2009. The "pooled" relatedness can be invalid and meaningless for the regression if the structure of task relatedness has large changes during the time period, and $\widetilde{\phi}_{i,x}$ is not fully exogenous since the dependent variable is determined by the changes in RCAs. The validity of the pooled task space is justified, since I find that structure of task space remains quite stable in the 15 years (see appendix for details). And to exclude the possibility of endogeneity, I derive the bilateral relatedness indices using only the RCAs from the first five years, i.e. 1995 to 1999. Then I construct the explanatory variables using these new relatedness indices $\phi^{95-99}_{y\to x}$, and together with the export structure and PRODY in 2000 to construct the proximity measure $\widetilde{\phi}^{*}_{i,x}$. In the appendix I replicate all regressions in table 4 and 5 using the new set of variables to analyse the development of new comparative advantages from 2000 to 2009 which do not have a time overlap with the new explanatory variables. Since the time period is shorter than from 1995 to 2009, the propensity in gaining a comparative advantage is smaller which makes the coefficients less significant and the marginal effects smaller compared with the baseline regressions. But the coefficient for $\widetilde{\phi}^*_{i,x}$ is positive in all cases, and in most cases also statistically significant.

7 Concluding Remarks and the Implication on Development Strategies

This paper investigates the relatedness between different tasks in a world where offshoring is pervasive and comparative advantage is realised at task level. Using the world input-output tables, I control for the trade in intermediate inputs and derive the value-added export (VAE) based revealed comparative advantages (RCAs) of tasks. I follow the methodology as in Hidalgo *et al.* (2007) and derive the task space, i.e. the bilateral relatedness between each pair of tasks based on the conditional probability that a country has comparative advantages in both tasks.

I find that in general tasks are more closely related with other tasks at the same skill level, rather than with the different skilled tasks in a same industry. This is especially the case for low-skilled tasks. The tight relatedness across low-skilled tasks in almost all industries indicates the horizontal upgrading, i.e. shift of low-skilled employment into new tasks with a higher economic potential, is likely to take place. On the other hand the relatedness between low- and higher-skilled tasks is rather low even for tasks within the same industry, with a few exceptions in traditional industries, utilities, and trade and logistics related services. It suggests that developing countries may face high difficulty in entering many medium- and high-skilled tasks; vertical upgrading is unlikely to take place especially in the GVCs of sophisticated products like electronics and automobile. I tested my task space against the actual structural change paths for 40 countries from 1995 to 2009 using a set of probit regressions. The econometric analyses shows that new comparative advantages are more likely to develop in the tasks that have strong proximity with the country's current comparative advantages. However, I find no support for vertical upgrading; in general having a comparative advantage in a low-skilled task does not increase the probability that higher-skilled tasks in the same industry will develop.

The finding in this paper has important implications for development strategy and is related with a long-going debate concerning whether a developing country should follow or defy its current comparative advantages in order to have successful structural change (see, e.g. Lin and Chang 2009). Lin and Monga (2011) argue that a developing country should not aim at too far-going targets on advanced activities in the beginning, but should instead stick to the current comparative advantages in low-skilled labour and try to find the industries where they have the "latent comparative advantage". They also argue that the government should not have strong distortive industrial policies, but instead play a role in information and infrastructure provision. This strategy is partially supported by the outcome from this paper, but my task space also shows several potential challenges for this strategy from a global value chain perspective.

My result supports the argument that in the early stage of development a country should not defy its current comparative advantages, namely the abundance of low-skilled labour. Low-skilled tasks in different industries show large hetergeneity in terms of their potentials for economic growth. Some low-skilled tasks, especially those in sophisticated products' GVCs, have a much higher economic potential than other traditional low-skilled tasks. PRODY of low-skilled tasks is around \$20,000 in automobile, machinery, and electronics industries compared with \$10,000 in textile and agriculture. Since all kinds of low-skilled tasks are closely related, the horizontal upgrading from traditional industries towards low-skilled tasks in sophisticated global value chains is therefore a strategy which is relatively easy to achieve and should bring much gain to the economy.

In the long-run the country must have vertical upgrading in order to achieve further development. Similar as Taglioni and Winkler (2016), I find the participation in GVC and the upgrading through GVC are two very different processes; and the later does not seem to be automatic. Participation in GVCs in general does not increase the probability of upgrading; a country might instead stuck in the low-skilled tasks due to the lack of relatedness with higher-skilled tasks in the value chain. It echoes the findings by Jarreau and Poncet (2012) and Lemoine and Ünal-Kesenci (2004) regarding the role of processing trade in China's economic growth. Jarreau and Poncet show that the regions that participate in the production of sophisticated goods indeed grow faster. However, the contribution to economic growth originates only from ordinary firms who produce and export sophisticated goods; foreign firms and processing trade firms have neither direct nor indirect effect on growth. Lemoine and Ünal-Kesenci focus on the electronics industry and they find that imported foreign technology has improved the competitiveness of processing exporters and helped China's export upgrading, but there is little technological diffusion that helps other Chinese firms, therefore the further gains from foreign technology are rather limited.

The difficulty in vertical upgrading suggests that a country must be careful in choosing which kind of value chain it is going to enter. Randomly participating in a value chain may generate a large gain in the country's economic potentials, but it is possibly once-off and there is no guarantee for further vertical upgrading. Therefore, before participating in a certain GVC, the policy makers should think whether the medium- and high-skilled tasks in that GVC will have high proximity with the economic structure of the country in the near future. This decision making is in analogy with the identification of latent comparative advantage as in Lin and Monga (2011), who suggest that the policy makers should look at a model country which is rapid growing, has around 100% higher GDP per capita, and had a similar socio-economic structure in the past. And then the latent comparative advantage should be found in a list of tradable products or services which the model country has produced for around 20 years. However, it is uncertain whether this imitating strategy remains valid under the recent wave of off-shoring and production fragmentation.

First, the organization of production has drastically changed in the past decades due to offshoring and technical change. The way a product is currently produced can be very different from 20 years ago; the production of a same product (or service) may entail completely different tasks which do not necessarily fit the current socio-economic conditions of the country. Second, my task space shows a same picture as in Baldwin (2013) that joining a value chain is a quicker way for industrialisation when offshoring is feasible, but it is questionable whether such "industrialisation" is meaningful in the long run. When the division of tasks in a value chain can be finely specialised and unbundled across borders, multinational companies minimize their production costs by re-organizing the production processes such that the offshored tasks are specially designed to fit the conditions in the host country. On one hand it means when a host country aims in entering a particular global value chain, it may always find some tasks that are specially designed for it, such that the country faces less obstacles in participating. However on the other hand, the name and feature of the products become less meaningful for the host country's future development. The offshored tasks are usually standardized due to the finer task specialisation which contains less product-specific core technologies, and can be similar in very different products. Tasks in more sophisticated products may end up in even less opportunities for learning which leads to a lower relatedness with the low-skilled tasks in the value chain. Sewing machine operators in a textile processing firm may gradually gain the experience on how Italians design their cloths, but the assemblers of smart phones will have nearly no chance nor capability to understand how Americans and Japanese make their chips.

Therefore, the policy makers should be more forward-looking and the policies should target on actual activities instead of products or industries. The development strategies should differ across different stages of development. Starting from a low level of development, it is worthwhile to follow the comparative advantage in low-skilled labour and try to achieve horizontal upgrading towards sophisticated value chains. But the exact direction should be carefully chosen in accordance with the country's own circumstances. And a different set of policies is needed when the country is at the stage for vertical upgrading, in which specific stimulation policies might still be necessary.

As a final remark, the paper has two noteworthy limitations. First, a task is defined as the activities in an industry that are performed by a particular skilled labour. This is an educational attainment based measure and it is possible that higher-skilled labour performs tasks that require lower skill levels. A better definition of task could be, for example, the exact occupations of workers. However the data on occupation structure by industry data are currently not available for many countries, so in this paper I stick to the educational attainment measure. Secondly, WIOD database covers a limited set of developing countries (Eastern European and the BRIC), and most of them are relatively successful in development in the past decades. It is therefore uncertain whether the structural of task space is the same for other developing countries at a lower level of development. The limited coverage for developing countries also force me to limit the scope of econometric tests in the direction of new comparative advantages, while a more interesting research question would be why some country successfully upgrade and why some not. To test the role of task relatedness in (overall) structural upgrading and economic growth, a much wider set of developing countries is needed, including both successful and stagnated ones.

References

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

Baldwin, Richard. 2006. "Globalisation: the great unbundling(s)." Economic Council of Finland.

———. 2013. "Trade and industrialization after globalization's second unbundling: how building and joining a supply chain are different and why it matters." In *Globalization in an age of crisis: Multilateral economic cooperation in the twenty-first century,* 165–212. University of Chicago Press.

Becker, Sascha O., Karolina Ekholm, and Marc-Andreas Muendler. 2013. "Offshoring and the onshore composition of tasks and skills." *Journal of International Economics* 90 (1): 91–106.

Bryce, David J., and Sidney G. Winter. 2009. "A general interindustry relatedness index." *Management Science* 55 (9): 1570–1585.

Clark, Colin. 1940. *The conditions of economic progress.* Alianza Editorial SA, Madrid.

Conley, Timothy G., and Bill Dupor. 2003. "A spatial analysis of sectoral complementarity." *Journal of Political Economy* 111 (2): 311–352.

Dedrick, Jason, Kenneth L. Kraemer, and Greg Linden. 2009. "Who profits from innovation in global value chains?: a study of the iPod and notebook PCs." *Industrial and Corporate Change:* 1–36.

Ellison, Glenn, Edward L. Glaeser, and William R. Kerr. 2010. "What causes industry agglomeration? Evidence from coagglomeration patterns." *American Economic Review* 100 (3): 1195–1213.

Grossman, Gene M., and Esteban Rossi-Hansberg. 2008. "Trading tasks: A simple theory of offshoring." *American Economic Review* 98 (5): 1978–1997.

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

Hausmann, Ricardo, and Dani Rodrik. 2003. "Economic development as self-discovery." *Journal of Development Economics* 72 (2): 603–633.

Hidalgo, César A., Bailey Klinger, Albert-Laszlo Barabási, and Ricardo Hausmann. 2005. "The product space and its consequences for economic growth." Unpublished working paper, the earlier version for the Science paper.

———. 2007. "The product space conditions the development of nations." *Science* 317 (5837): 482–487.

Hummels, David, Jun Ishii, and Kei-Mu Yi. 2001. "The nature and growth of vertical specialization in world trade." *Journal of international Economics* 54 (1): 75–96.

Jarreau, Joachim, and Sandra Poncet. 2012. "Export sophistication and economic growth: Evidence from China." *Journal of Development Economics* 97 (2): 281–292.

Johnson, Robert C., and Guillermo Noguera. 2012. "Accounting for intermediates: Production sharing and trade in value added." *Journal of International Economics* 86 (2): 224–236.

Kali, Raja, Javier Reyes, Joshua McGee, and Stuart Shirrell. 2013. "Growth networks." *Journal of Development Economics* 101:216–227.

Koopman, Robert, Zhi Wang, and Shangjin Wei. 2012. "Estimating domestic content in exports when processing trade is pervasive." *Journal of development economics* 99 (1): 178–189.

———. 2014. "Tracing value-added and double counting in gross exports." *American Economic Review* 104 (2): 459–494.

Lall, Sanjaya, John Weiss, and Jinkang Zhang. 2006. "The "sophistication" of exports: a new trade measure." *World Development* 34 (2): 222–237.

Lemoine, Francoise, and Deniz Ünal-Kesenci. 2004. "Assembly trade and technology transfer: the case of China." *World development* 32 (5): 829–850.

Leontief, Wassily. 1949. "Structural matrices of national economies." *Econometrica: Journal of the Econometric Society:* 273–282.

Lewis, W. Arthur. 1954. "Economic development with unlimited supplies of labour." *The Manchester School* 22 (2): 139–191.

Lin, Justin Yifu. 2012. *New structural economics: A framework for rethinking development and policy.* World Bank Publications.

Lin, Justin Yifu, and Ha-Joon Chang. 2009. "Should industrial policy in developing countries conform to comparative advantage or defy it? A debate between Justin Lin and Ha-Joon Chang." *Development Policy Review* 27 (5): 483–502.

Lin, Justin Yifu, and Yan Wang. 2008. "China's integration with the world: Development as a process of learning and industrial upgrading." *World Bank Policy Research Working Paper Series.*

Lin, Justin, and Célestin Monga. 2011. "DPR debate: Growth identification and facilitation: The role of the state in the dynamics of structural change." *Development Policy Review* 29 (3): 259–310.

Los, Bart, Marcel P. Timmer, and Gaaitzen J. de Vries. 2015. "How important are exports for job growth in China? A demand side analysis." *Journal of Comparative Economics* 43 (1): 19–32.

———. 2016. "Tracing Value-Added and Double Counting in Gross Exports: Comment." *American Economic Review:* 1958–1966.

Marshall, Alfred. 1920. *Principles of Economics.* Macmillan, London.

Nedelkoska, Ljubica, Frank Neffke, and Simon Wiederhold. 2015. “Skill mismatch and the costs of job displacement.” Annual Meeting of the American Economic Association.

Neffke, Frank, and Martin Henning. 2008. “Revealed relatedness: Mapping industry space.” Papers in Evolutionary Economic Geography, University of Utrecht 8:19.

———. 2013. “Skill relatedness and firm diversification.” *Strategic Management Journal* 34 (3): 297–316.

Neffke, Frank, Martin Henning, and Ron Boschma. 2011. “How do regions diversify over time? Industry relatedness and the development of new growth paths in regions.” *Economic Geography* 87 (3): 237–265.

Taglioni, Daria, and Deborah Winkler. 2016. *Making Global Value Chains Work for Development.* World Bank, Washington DC.

Teece, David J., Richard Rumelt, Giovanni Dosi, and Sidney G. Winter. 1994. “Understanding corporate coherence: Theory and evidence.” *Journal of Economic Behavior & Organization* 23 (1): 1–30.

Timmer, Marcel P., Bart Los Abdul A. Erumban, Robert Stehrer, and Gaaitzen J. de Vries. 2014. “Slicing up global value chains.” *Journal of Economic Perspectives* 28 (2): 99–118.

Timmer, Marcel P., Erik Dietzenbacher, Bart Los, Robert Stehrer, and Gaaitzen J. de Vries. 2015. “An illustrated user guide to the world input–output database: the case of global automotive production.” *Review of International Economics* 23 (3): 575–605.

APPENDIX

Table 6: Industry and Activitiy Code List

Skill Level / ISIC.Rev3 Sector	Low	Medium	High
AtB: Agriculture, Hunting, Forestry and Fishing	1	29	57
C: Mining and Quarrying	2	30	58
15t16: Food, Beverages and Tobacco	3	31	59
17t18: Textiles and Textile Products	4	32	60
19: Leather, Leather Products and Footwear	5	33	61
20: Wood and Products of Wood and Cork	6	34	62
21t22: Pulp, Paper, Printing and Publishing	7	35	63
23: Coke, Refined Petroleum and Nuclear Fuel	8	36	64
24: Chemicals and Chemical Products	9	37	65
25: Rubber and Plastic	10	38	66
26: Other Non-Metallic Mineral	11	39	67
27t28: Basic Metals and Fabricated Metal	12	40	68
29: Machinery, Not elsewhere classified	13	41	69
30t33: Electrical and Optical Equipment	14	42	70
34t35: Transport Equipment	15	43	71
36t37: Manufacturing, Not elsewhere classified; Recycling	16	44	72
E: Electricity, Gas and Water Supply	17	45	73
F: Construction	18	46	74
50: Sale and Repair of Motor Vehicles and Motorcycles; Retail Sale of Fuel	19	47	75
51: Wholesale Trade, Except of Motor Vehicles and Motorcycles	20	48	76
52: Retail Trade and Repair, Except of Motor Vehicles and Motorcycles	21	49	77
60: Inland Transport	22	50	78
61: Water Transport	23	51	79
62: Air Transport	24	52	80
64: Post and Telecommunication	25	53	81
J: Financial Intermediation	26	54	82
70: Real Estate Activities	27	55	83
71t74: Renting of Machinery & Equipment and Other Business Services	28	56	84

Table 7: Chinese Employment in Each Task, 1995 and 2009 (billion working hours)

	1995				2009				Chg. in
	L	M	H	%MH	L	M	H	%MH	%MH
Agriculture	488.3	18.16	0.28	3.64	442.8	26.53	0.12	5.68	2.04
Mining	16.04	9.91	0.28	38.85	13.71	11.45	0.94	47.48	8.63
Food Manufacturing	13.28	8.61	0.39	40.39	18.77	14.24	1.40	45.47	5.08
Textile	25.41	9.69	0.20	28.02	34.43	15.38	0.70	31.83	3.81
Leather Products	4.38	1.34	0.02	23.68	10.72	3.82	0.15	27.07	3.39
Furniture	5.11	1.92	0.05	27.82	12.89	5.68	0.30	31.68	3.87
Paper and Publishing	5.77	3.53	0.11	38.66	13.33	9.55	0.62	43.28	4.63
Refinery	0.95	1.02	0.07	53.16	1.12	1.39	0.20	58.81	5.64
Chemistry	8.51	7.73	0.55	49.28	11.39	12.11	1.87	55.09	5.81
Rubber and Plastic	6.70	3.43	0.11	34.60	15.03	9.02	0.64	39.12	4.52
Non-metal Minerals	21.13	8.02	0.21	28.02	16.23	7.21	0.41	31.96	3.93
Metal	13.64	9.16	0.45	41.33	15.24	11.98	1.28	46.54	5.21
Machinary	13.05	9.91	0.60	44.62	16.03	14.25	1.89	50.19	5.57
Electronics	8.67	7.56	0.63	48.57	21.78	22.24	4.02	54.67	6.10
Automobile	5.34	5.20	0.35	50.95	6.86	7.81	1.16	56.68	5.73
Other Manufacturing	13.69	5.12	0.14	27.76	12.26	5.36	0.33	31.72	3.96
Utility	2.24	3.61	0.35	63.91	2.25	6.12	1.58	77.37	13.45
Construction	50.44	28.05	1.25	36.74	75.45	51.45	4.17	42.44	5.69
Sale of Automobile	n.a.	n.a.	n.a.	n.a	n.a.	n.a.	n.a.	n.a.	n.a.
Wholesaling	8.06	13.57	1.41	65.03	6.64	18.84	5.53	78.59	13.56
Retailing	21.80	38.16	2.01	64.82	19.01	56.05	8.35	77.21	12.39
Inland Transport	14.10	14.82	0.49	52.05	20.38	24.28	1.94	56.27	4.21
Water Transport	1.91	1.54	0.13	46.66	2.40	2.20	0.45	52.40	5.74
Air Transport	0.20	0.77	0.19	82.54	0.32	1.36	0.83	87.34	4.80
Post and Communication	1.32	5.88	1.51	84.89	1.57	7.95	4.96	89.15	4.27
Finance	0.13	5.40	0.58	97.84	0.18	7.14	2.33	98.12	0.28
Real Estate	0.88	0.96	0.12	55.34	1.72	1.48	0.31	51.04	-4.29
Business Services	3.10	3.67	0.99	60.05	2.90	2.95	1.86	62.36	2.31
All	754.18	226.74	13.47	24.16	795.39	357.86	48.35	33.81	9.65

Source: Socio-economic account in the WIOD dataset. L, M, H stands for the low-, medium- and high-skilled tasks in each industry. Columns "%MH" are the shares of medium- and high-skill labour in total industrial employment. The last column compares the changes in medium- and high-skilled labour share between 1995 and 2009.

The Stability of Task Space Across Different Time Periods

The relatedness indices can be derived separately for each year in principle, but in this paper I follow Hidalgo *et al.* (2007) and use a "pooled" relatedness by taking the average of relatedness indices across 1995 to 2009. To justify this choice, I first explore the stability of my task space. I divide the 15 years into three five-year periods, and derive three sets of task spaces based on the data within each period. Table 8 reports the correlations between task spaces of different years; the upper panel reports Pearson correlation (i.e. the conventional correlation) and the lower panel reports the Spearman's rank-order correlation. The task space is changing but the structure is relatively consistent over time; correlation between two consecutive five-year periods is higher than 0.9, which is a bit higher than the correlation of product space reported in Hidalgo *et al.* (2007) and Neffke *et al.* (2008). The correlation between a single period's task space and the pooled task space for the whole period is also around 0.9. Given this high stability of task space, it is therefore safe to assume the structure of relatedness between tasks stays constant over the whole period, represented by the pooled task space using data of all years.

Table 8: Stablity of the Task Space across Different Years

Pearson Correlation	95-99	00-04	04-09	95-09
95-99	1			0.7998
00-04	0.9263	1		0.8295
04-09	0.8407	0.9173	1	0.8074

Rank Order Corr.	95-99	00-04	04-09	95-09
95-99	1			0.8801
00-04	0.9190	1		0.9104
04-09	0.8273	0.9099	1	0.8804

Table 9: Regression Results (Robustness Checks)

I - Main Results

	Baseline		Use $\phi_x^{\max}$	Both	Symmetric Relatedness		Developing Econ	
	(1)	MEF	(1a)	(1b)	(2)	MEF	(3)	MEF
$\widetilde{\phi}^*_{x,i}$	2.350	0.075		2.950	3.153	0.099	1.439	0.067
	0.783^{**}	0.028^{**}		1.006^{**}	0.932^{***}	0.034^{**}	0.939	0.045
$\phi^{\max}_{x,i}$			0.519	-0.621				
			0.508	0.652				
Potential$_{00}$	1.083	0.035	1.095	1.014	1.051	0.033	0.757	0.035
	0.239^{***}	0.009^{***}	0.245^{***}	0.248^{***}	0.238^{***}	0.008^{***}	0.278^{**}	0.013^{**}
RCA$_{00}$	3.321	0.106	3.295	3.410	3.336	0.104	3.735	0.173
	0.417^{***}	0.019^{***}	0.423^{***}	0.429^{***}	0.418^{***}	0.019^{***}	0.521^{***}	0.034^{***}
Const.	-3.739		-3.630	-3.498	-3.833		-3.602	
	0.281^{***}		0.370^{***}	0.373^{***}	0.290^{***}		0.340^{***}	
R^2	0.222		0.208	0.223	0.225		0.242	
Obs.	1584		1584	1584	1584		814	
# Positive	68	(4.3%)	68	68	68		46	(5.7%)

II - Testing Vertical Upgrading

	M&H Tasks, All Observations				M&H Tasks, Developing Countries			
	(4a)	(4b)	(4c)	(4d)	(5a)	(5b)	(5c)	(5d)
$\widetilde{\phi}_{x,i}$	2.486	2.326	2.509		0.958	0.506	1.204	
	0.985^{*}	1.009^{*}	0.989^{**}		1.236	1.324	1.273	
Potential$_{00}$	1.800	1.673	1.847	1.661	1.409	1.169	1.601	1.465
	0.400^{***}	0.430^{***}	0.428^{***}	0.419^{***}	0.556^{*}	0.597^{*}	0.610^{**}	0.594^{*}
RCA$_{00}$	3.611	3.538	3.639	3.665	4.027	3.930	4.096	4.102
	0.508^{***}	0.516^{***}	0.517^{***}	0.514^{***}	0.653^{***}	0.660^{***}	0.663^{***}	0.661^{***}
RCA$_L$		0.041				0.065		
		0.052				0.061		
IsRCA$_L$			-0.049	-0.021			-0.178	-0.128
			0.163	0.162			0.224	0.218
Const.	-4.065	-4.035	-4.070	-3.634	-3.830	-3.739	-3.887	3.663
	0.368^{***}	0.369^{***}	0.369^{***}	0.313^{***}	0.482^{***}	0.487^{***}	0.492^{***}	0.422^{***}
R^2	0.215	0.216	0.216	0.200	0.229	0.232	0.231	0.227
Obs.	1025				512			
# Positive	53	(5.2%)			35	(6.8%)		

Notes: Replication of the specifications in Table 4. The proximity indicator $\widetilde{\phi}^*_{x,i}$ is constructed from the relatedness based on the year 1995 to 1999, and the economic structure in year 2000. The dependent variable is a dummy indicating whether a task with RCA lower than 0.75 in 2000 has gained a comparative advantage in 2009. The dependent variables is a dummy indicating whether a task with RCA lower than 0.75 in 2000 has gained a comparative advantage in 2009.

Table 10: Regression Results for Alternative Specifications (Robustness Check)

	All $RCA_{00} < 1$		RCA Growth> 0.25		Changes in RCA	
	(6)	MEF	(7)	MEF	(8)	(9)
$\widetilde{\phi}^*_{x,i}$	1.851	0.119	2.370	0.387	0.134	0.217
	0.584^{**}	0.039^{**}	0.521^{***}	0.085^{***}	0.065^{**}	0.181
Potential_{00}	1.302	0.084	1.415	0.231	0.179	0.557
	0.189^{***}	0.013^{***}	0.160^{***}	0.022^{***}	0.012^{***}	0.034^{***}
RCA_{00}	3.169	0.204	1.308	0.214	-0.053	-0.413
	0.225^{***}	0.022^{***}	0.213^{***}	0.034^{***}	0.023^{*}	0.065^{***}
Const.	-3.598		-2.091		0.051	0.181
	0.198^{***}		0.135^{***}		0.013^{***}	0.037^{***}
R^2	0.290		0.126		0.122	0.165
Obs.	1990		1584		1570	1529
# Positive	190	(9.6%)	208	(13.2%)	-	-

Notes: Replication of specifications in table 5. The proximity indicator $\widetilde{\phi}^*_{x,i}$ is constructed from the relatedness based on the year 1995 to 1999, and the economic structure in year 2000. The dependent variable is similar as in table 5 but based on the changes between 2009 and 2000 (instead of 1995).