博士論文（要約）

Skill Distribution, Firm Organization, and Comparative Advantage: From a Production Linkage Perspective

（スキル分布・企業組織・比較優位：
生産リンケージの視点から）

明日山　陽子

Ackowledgements

First, I would like to express my sincere gratitude to my dissertation committee, Professors Yuji Genda, Hideo Owan, Ryuichi Tanaka, Ayako Kondo, and Yasuyuki Todo. I especially thank my main adviser, Professor Yuji Genda, for his many insightful comments and continuous guidance on my work. I also especially thank Professor Hideo Owan for his valuable comments on Chapter 4, and Professor Ryuichi Tanaka who kindly reviewed Chapter 2 as a discussant at the 2016 Spring Meeting of the Japanese Economic Association.

I am also grateful to my employer, Japan External Trade Organization (JETRO), and its affiliated research institution, Institute of Developing Economies (IDE). Previous versions of all three empirical chapters in this dissertation were released as the IDE discussion papers. In addition, previous versions of Chapters 2 and 3 were written as the research output of the IDE's research projects. JETRO and IDE have continuously supported my work by providing financial assistance and ideal research environment. In particular, I would like to thank Dr. Tomohiro Machikita, who gave me various valuable comments on Chapters 2 and 4, and Drs. Tatsufumi Yamagata, Momoko Kawakami, and Mariko Watanabe, who supported me when I applied for the PhD program.

I am also indebted to many people: participants at the Labor Economics Seminar for the Graduate Students at the University of Tokyo, those attending at Professor Genda's classes, and my colleagues at the IDE. Some of them are in the acknowledgement section of each chapter.

Finally, I greatly thank my husband, daughter, and parents for their continuous supports. Without their supports, my study at the PhD program would not have been feasible, and this dissertation would not have been written.

Contents

Chapter 1

Introduction

1. Overview of the Dissertation

This dissertation consists of three empirical research papers (Chapters 2, 3, and 4) and one concluding chapter (Chapter 5). All three empirical chapters examine how the differences in technological characteristics across industries influence the economic development patterns. More precisely, they empirically show that the impact of human/social capitals on a certain outcome (wages, export, and delegation to workers, depending on the chapter) depends on the technological characteristics of the industry (length of production chains or coordination needs). Such different impacts across industries in turn shape the patterns of economic development.

Using Indian data from 1999 and 2009, Chapter 2 empirically shows that high-skilled workers are sorted into industries with shorter domestic production chains, where wage returns to skill are higher. This skill-sorting pattern, which I call "negative sorting," contradicts the majority of theoretical studies, which predict that high-skilled individuals work in sectors with longer production chains ("positive sorting"). I hypothesize that negative sorting occurs in India because of substantial quality deterioration along the production chains as a result of a large pool of low-skilled labor, poor infrastructure, and less-advanced technology. When quality deterioration concomitant with an increase in production chain length is substantial, the wages of high-skilled workers are also dragged down substantially in industries with longer production chains. Consequently, high-skilled workers choose industries with shorter production chains.

By using industry export and skill distribution data on 58 economies around the world in 2000, Chapter 3 empirically shows that a country with a greater (respectively, lower) dispersion of skills exports relatively more in industries with shorter (longer) production chains. Based on the skill-sorting mechanism outlined in Chapter 2, Chapter 3 hypothesizes that in countries with greater skill dispersion, the degree of quality deterioration along production chains is more substantial. As a result, negative sorting is more likely to occur and the skill-intensity gap between sectors increases. Provided that *ceteris paribus*, industry productivity is determined by its skill intensity, greater skill dispersion leads to a comparative advantage in industries with shorter production chains.

Chapter 4 examines worker-level data from 14 countries and empirically explains why the degree of delegating authority to non-managerial and non-supervisory workers substantially varies across countries and industries. I focus on the differences in the following two types of delegation costs: the region-specific social capital that proxies workers' degree of self-centeredness and the industry-specific need for coordination. The empirical results show that the negative association between coordination needs and decentralization is mitigated in regions with higher social capital (i.e., lower self-centeredness of workers). In particular, when the self-centeredness of workers (respectively, need for coordination) is very low, the degree of delegation is always high regardless of the level of the need for coordination (self-centeredness of workers). I also find positive associations between delegation and its benefits, including job satisfaction, wages (proxy for productivity), and skill upgrading of workers.

Finally, Chapter 5 addresses several areas for future research.

2. Comparison of the Three Empirical Chapters

Table 2.1 compares the three empirical chapters, which share certain common features. First, as mentioned above, all chapters show that the impact of human/social capitals on a certain outcome depends on the technological characteristics of a given industry. Chapter 2 empirically shows that returns to skill (impact of skill on wages) are higher in industries with shorter domestic production chains in India. Chapter 3 presents that the impact of a country's skill dispersion on export is more positive in industries with shorter domestic production chains. Finally, Chapter 4 shows that the impact of regional social capital on delegation is more positive in industries with greater coordination needs.

Second, these different impacts across industries shape the patterns of economic development. In Chapter 2, inter-industry skill wage differentials affect the skill allocation across industries. The negative (skill) sorting in India is likely to promote the development of industries with shorter production chains, but impede that of industries with longer production chains, which generally have larger production and employment spillover. Chapter 3 directly shows that a country with a greater (respectively, lower) dispersion of skills has a comparative advantage in industries with shorter (longer) production chains. Chapter 4 also implies that regions with higher social capital (lower self-centeredness of workers) may have a comparative advantage in industries requiring greater coordination and exhibit lower inequalities in terms of both income and skill.

Third, the primary estimating equation in each chapter takes the following form:

$$Y_{ij} = \beta_1 Z_{ij} + \beta_2 X_i + \beta_3 Z_{ij} * X_i + \gamma C_{ij} + F_j + \varepsilon_{ij}\text{, or}$$

$$Y_{ij} = \beta_1 Z_{ij} + \beta_3 Z_{ij} * X_i + \gamma C_{ij} + F_j + F_i + \varepsilon_{ij}, \qquad (2.1)$$

where subscript i indicates worker (Chapter 2), exporting country (Chapter 3), or region (Chapter 4); j indicates industry; Y_{ij} is the dependent variable, which is either wage (Chapter 2), export (Chapter 3), or delegation (Chapter 4); X_i is i's characteristics, which is skill level (Chapter 2), skill dispersion (Chapter 3), and social capital (Chapter 4); Z_{ij} is an industry's technological characteristics, which is the length of domestic production chains (Chapters 2 and 3) or the need for coordination (Chapter 4); C_{ij} is the other control variables including a constant; F_j and F_i are industry dummies and i dummies, respectively; and ε_{ij} is the error term. The different impact of X_i on Y_{ij} across industries (or different impact of Z_{ij} on Y_{ij}, which depends on X_i) is captured by the coefficient β_3. This empirical strategy, which examines the coefficient of the interaction term between X_i and Z_{ij}, is common in studies in macroeconomics and international trade (see the next subsection and Ciccone and Papaioannou 2016).

Fourth, in all three chapters, I use the column sum of the Leontief inverse coefficient of industry *j*, which is computed from the input–output tables, as the primary measure for the industry's technological characteristics, Z_{ij}. Denote this measure $Leon_j$. This $Leon_j$ measures the amount of intermediate inputs that industry *j* requires, both directly and indirectly, to produce one dollar's worth of output. It measures the scope of production linkages with intermediate input industries. Fally (2012) also considers that $Leon_j$ measures "the number of production stages embodied in each product" (Fally 2012: 2). Regardless of countries, $Leon_j$ tends to be larger in many manufacturing industries, in particular transport equipment and basic metals, whereas it tends to be smaller in most primary and service industries. In Chapters 2 and 3, $Leon_j$, which is computed based on domestic intermediate inputs, is used as a measure for the length of domestic production chains. In these chapters, it is assumed that as the amount of necessary intermediate inputs increases (i.e., as the length of production chains increases), the quality of these inputs when combined degrades more, similar to Kremer's (1993) O-ring theory. For example, if the probability of malfunction in each part is 1%, then that of a product composed of two units of the part becomes 1.99% (= 1-0.99*0.99)*100). In Chapter 4, $Leon_j$, which is computed based on both domestic and imported intermediate inputs is used as a proxy for an industry's coordination needs. To produce a perfect-quality final product, each intermediate input needs to be of good quality and fit each other, and cost-effective by utilizing economies of scale. Thus, coordination becomes more intense and important for a firm's profit with an increase in the amount of intermediate inputs.

Fifth, all three chapters cover not only the manufacturing sector but also the primary (Chapters 2, 3, and 4) and service sectors (Chapters 2 and 4). Restricting the sample to manufacturing may have an advantage in reducing unobservable factors. On the other hand, such a restriction is not favorable when analyzing the individual choice of an industry or a country's comparative advantage. Furthermore, the border of manufacturing and services is becoming less clear (Kastalli and Looy 2013).

Finally, all three chapters check the robustness of the results by examining multiple indices and correcting for the possible selection or other endogeneity biases.

3. Contribution of the Dissertation

This dissertation mainly contributes to the literature by empirically showing that the effect of human/social capitals on a certain outcome—wages, export, and delegation, or more broadly, economic development—is not homogeneous: it varies depending on a given industry's scope of production linkages.

The positive effect of human capital on individual earnings has been continuously found in numerous empirical studies since the pioneering works on human capital (Becker 1964; Mincer 1974).[1] The importance of human capital on a country's economic development (cross-country differences in long-term growth rate or per capita income) has also been continuously emphasized since the early empirical studies, such as Barro (1991) and Mankiw *et al*. (1992).[2] Although later empirical studies (e.g., Klenow and Rodríguez-Clare 1997; Hall and Jones 1999) find a modest role of human

[1] See also the surveys conducted by Willis (1986) and Card (1999).

[2] Much earlier studies, such as Griliches (1970), also emphasize the importance of human capital in the economic development of the US.

capital and emphasize a significant role of total factor productivity, recent studies (e.g., Jones 2014; Hendricks and Shoellman 2016), which try to measure human capital more precisely, find a much larger effect of human capital on cross-country income differences. A limited number of studies also evaluate the impact of the distribution of human capital, and generally find its negative association with a country's economic growth (see the survey conducted by Sauer and Zagler 2012).

The importance of social capital for successful economic development was first emphasized in sociology and political science (Banfield 1958; Coleman 1990; Putnam *et al*. 1993; Fukuyama 1995; Putnam 2000). Since the mid-1990s, economists also began conducting empirical investigations on the effect of social capital on economic outcome, such as economic growth (Knack and Keefer 1997; Algan and Cahuc 2010), financial development (Guiso *et al*. 2004), and the organization and size of firms (Bloom *et al*. 2012; Cingano and Pinotti 2016).[3] In most economic studies, social capital is measured by the level of trust.

Since Rajan and Zingales (1998), many empirical studies in macroeconomics and international trade have shown that the effect of a country (or region) characteristic X on a certain economic outcome Y (e.g., economic growth, comparative advantage) depends on a certain industry characteristic Z. As Ciccone and Papaioannou (2016) noted, these studies generally estimate equations similar to equation (2.1). However, only a few studies characterize a country/region by the distribution of skill (Bombardini *et al*. 2012) or the level of social capital (Cingano and Pinotti 2016). Industries are characterized by their dependence on external sources of finance, capital or skill intensity, R&D intensity, and so on (see Appendix Table 1 of Cingano and Pinotti 2016).

[3] Also see the literature review conducted by Durlauf and Fafchamps (2005), Guiso *et al*. (2011), and Algan and Cahuc (2014). Economists have also tried to define social capital more precisely. For example, Guiso *et al*. (2011) define social capital as "those persistent and shared beliefs and values that help a group overcome the free rider problem in the pursuit of socially valuable activities" (Guiso *et al*. 2011: 422–423).

No studies characterize industries by the length of production chains, although similar flavors can be found in the skill substitutability measure in Bombardini *et al*. (2012) and the product complexity measure in Levchenko (2007).

Production chains have been analyzed in several contexts. First, several studies examine allocation patterns of workers (Kremer 1993), countries (Kremer 1993; Levine 2012; Costinot *et al*. 2013), and ownership rights (i.e., make-or-buy [outsourcing] decision) (Grossman and Rossi-Hansberg 2008; Antras and Chor 2013) along the domestic or global production chains. Second, complementarity and fragility (Kremer 1993; Blanchard and Kremer 1997; Jones 2010; Levine 2012) and the multiplier effect (Jones 2010) of the production chains are emphasized. These studies imply that because of complementarity or the multiplier effect, a small mistake in the production chain significantly reduces the output. This may in turn explain the cross-country income gaps that are far greater than the observed differences in factor endowments. Third, in some studies, the effect of the production chain length, which proxies the number of relation-specific suppliers, is analyzed under incomplete contracts (Blanchard and Kremer 1997; Levchenko 2007; Nunn 2007). The implication of these studies is that countries with poorer contract enforcement institutions suffer more from the inefficiency (i.e., underinvestment) in sectors with more relation-specific suppliers.

The focus of my analyses is inter-industry variation in the length of production chains (i.e., the scope of production linkages). In Chapters 2 and 3, I emphasize the mechanism that the aggregate quality of intermediate inputs deteriorates as the length of production chains increases. In Chapter 4, I focus on the aspect that greater production linkages require more coordination. Chapter 2 can be classified as the first strand of the above literature, but I find a skill sorting pattern of workers opposite of the one predicted by Kremer (1993).

In summary, this dissertation aims to add value to the literature by finding new nexuses among human/social capitals, industry technology, and economic development.

References

Algan, Yann, and Pierre Cahuc. 2010. "Inherited Trust and Growth." *American Economic Review*, 100(5): 2060-2092.

Algan, Yann, and Pierre Cahuc. 2014. "Trust, Growth, and Well-Being: New Evidence and Policy Implications." In Philippe Aghion and Steven N. Durlauf eds., *Handbook of Economic Growth*, Volume 2A.North-Holland.Chapter 2, 49-120.

Antràs, Pol, and Davin Chor. 2013. "Organizing the Global Value Chains." *Econometrica*, 81(6): 2127–2204.

Banfield, Edward C. 1958. *The Moral Basis of a Backward Society*. New York: Free Press.

Barro, Robert J. 1991. "Economic Growth in a Cross Section of Countries" *Quarterly Journal of Economics*, 106(2): 407-443.

Becker, Gary S. 1964. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. New York: Columbia University Press.

Blanchard, Olivier, and Michael Kremer. 1997. "Disorganization." *Quarterly Journal of Economics*, 112(4): 1091-1126.

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen. 2012. "The Organization of Firms across Countries." *Quarterly Journal of Economics*, 127(4): 1663-1705.

Bombardini, Matilde, Giovanni Gallipoli, and Germán Pupato. 2012. "Skill Dispersion and Trade Flows." *American Economic Review*, 102(5): 2327–2348.

Card, David. 1999. "The Causal Effect of Education on Earnings." In Orley C. Ashenfelter and David Card eds., *Handbook of Labor Economics*, Volume 3. Elsevier Science B.V. Chapter 30, 1801-1863.

Ciccone, Antonio, and Elias Papaioannou. 2016. "Estimating Cross-Industry Cross-Country Interaction Models Using Benchmark Industry Characteristics."

NBER Working Paper 22368.

Cingano, Federico, and Paolo Pinotti. 2016. “Trust, Firm Organization, and the Pattern of Comparative Advantage.” *Journal of International Economics*, 100: 1-13.

Coleman, James. 1988. “Social Capital in the Creation of Human Capital.” *American Journal of Sociology*, 94: 95-120.

Costinot, Arnaud, Jonathan Vogel, and Su Wang. 2013. "An Elementary Theory of Global Supply Chains." *Review of Economic Studies*, 80(1): 109-144.

Durlauf, Steven N., and Marcel Fafchamps. 2005. “Social Capital.” In Philippe Aghion and Steven N. Durlauf eds., *Handbook of Economic Growth*, Volume 1B. Elsevier B.V. Chapter 26, 1639-1699.

Fally, Thibault. 2012. “Production Staging: Measurement and Facts.” Unpublished working paper, August 2012, University of Colorado-Boulder.

Fukuyama, Francis. 1995. *Trust: The Social Virtues and the Creation of Prosperity*. New York: Free Press.

Griliches, Zvi. 1970. “Notes on the Role of Education in Production Functions and Growth Accounting.” In W. Lee Hansen, ed., *Education, Income, and Human Capital*. NBER, 71-127.

Grossman, Gene M., and Esteban Rossi-Hansberg. 2008. “Trading Tasks: A Simple Model of Off-Shoring.” *American Economic Review*, 98(5): 1978-1997.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2004. “The Role of Social Capital in Financial Development.” *America Economic Review*, 94(3): 526-556.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2011. “Civic Capital as the Missing Link” In Jess Benhabib, Alberto Bisin, and Matthew O. Jackson, eds. *Handbook of Social Economics*, Volume 1. North-Holland, Chapter 10, 417-480.

Hall, Robert E., and Charles I. Jones. 1999. “Why Do Some Countries Produce So Much More Output Per Worker Than Others?” *Quarterly Journal of Economics*, 114(1): 83–116.

Hendricks, Lutz, and Todd Schoellman. 2016. “Human Capital and Development Accounting: New Evidence from Wage Gains at Migration.” Unpublished working paper, March 2016.

Jones, Benjamin F., 2014. “The Human Capital Stock: A Generalized Approach” *American Economic Review*, 104(11): 3752-3777.

Jones, Charles I. 2011. “Intermediate Goods and Weak Links in the Theory of Economic Development.” *American Economic Journal: Macroeconomics*, 3(2): 1-28.

Kastalli, Ivanka Visnjic, and Bart Van Looy. 2013. “Servitization: Disentangling the Impact of Service Business Model Innovation on Manufacturing Firm Performance.” *Journal of Operations Management*, 31(4): 169-180.

Klenow, Peter J., and Andrés Rodríguez-Clare. 1997. “The Neoclassical Revival in Growth Economics: Has It Gone Too Far?” In Ben S. Bernanke and Julio Rotemberg eds., *NBER Macroeconomics Annual*, Vol. 12. Cambridge, MA: MIT Press: 73–103.

Knack, Stephen, and Philip Keefer. 1997. “Does Social Capital Have an Economic Payoff? A Cross-Country Investigation.” *Quarterly Journal of Economics*, 112(4): 1251–88.

Kremer, Michael. 1993. “The O-Ring Theory of Economic Development.” *Quarterly Journal of Economics*, 108(3): 551-575.

Levchenko, Andrei A. 2007. “Institutional Quality and International Trade.” *Review of Economic Studies*, 74(3): 791–819.

Levine, David K. 2012. “Production Chains.” *Review of Economic Dynamics*, 15(3): 271-282.

Mankiw, N. Gregory, David Romer, and David N. Weil. 1992. “A Contribution to the Empirics of Economic Growth.” *Quarterly Journal of Economics*, 107 (2): 407–37.

Mincer, Jacob. 1974. *Schooling, Experience and Earnings*. New York: Columbia

University Press.

Nunn, Nathan. 2007. "Relationship-Specificity, Incomplete Contracts, and the Pattern of Trade." *Quarterly Journal of Economics*, 122(2): 569-600.

Putnam, Robert D. 2000. *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon and Schuster.

Putnam, Robert D., Robert Leonardi, and Raffaella Y. Nanetti. 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, NJ: Princeton University Press.

Rajan, Raghuram G., and Luigi Zingales. 1998. "Financial Dependence and Growth." *American Economic Review*, 88(3): 559-586.

Sauer, Petra, and Martin Zagler. 2012. "Economic Growth and the Quantity and Distribution of Education: A Survey." *Journal of Economic Surveys*, 26(5): 933-951.

Willis, Robert J. 1986. "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions." In Orley C. Ashenfelter and R. Layard eds., *Handbook of Labor Economics*, Volume 1. Elsevier Science Publishers B.V., Chapter 10, 525-602.

Table 2.1 Comparison of the Three Chapters

	Chapter 2	Chapter 3	Chapter 4
Country coverage	India	58 developed and developing countries	14 (mostly OECD member) countries
Year coverage	1999, 2009	2000	2011–2012
Industry coverage	All sectors (57 industries)	Nonservice sector (18 industries)	All sectors (34 or 71 industries)
Common specification	$Y_{ij} = \beta_1 Z_{ij} + \beta_2 X_i + \beta_3 Z_{ij} * X_i + \gamma C_i + F_j + \varepsilon_{ij}$, or $Y_{ij} = \beta_1 Z_{ij} + \beta_3 Z_{ij} * X_i + \gamma C_i + F_j + F_i + \varepsilon_{ij}$, Where *i*: individual, region, or country; *j*: industry; *C*: control variables including a constant; *F*: fixed effects; ε : error term; and *Y*, *X*, and *Z*: see below.		
Y_{ij} : Dependent variable	**Wage** of a regular wage/salaried male employee	**Industry export** from country *x* to country *m*	**Degree of delegation** of a non-managerial/non-supervisory worker
X_i : *i*'s characteristics	**Worker's skill level** (4 indices)	**Exporting country's skill dispersion** (3 indices)	**Region's social capital** (which is higher as the **self-centeredness of workers** is lower) (6 indices)
Z_{ij} (Z_j): Industry characteristics	**Length of domestic production chains** = Column sum of the Leontief inverse coefficient based on *domestic* intermediate inputs	**Length of domestic production chains** = Column sum of the Leontief inverse coefficient based on *domestic* intermediate inputs	**Need for coordination** = 1) Column sum of the Leontief inverse coefficient based on *total* intermediate inputs; or 2) skill complementarity measure developed by Bombardini *et al*. (2012)
Results	Return to skill is higher in industries with shorter production chains ($\beta_3 < 0$) in India. As a result, high-skilled workers are sorted into these industries.	Countries with a greater dispersion of skills export relatively more in industries with shorter production chains ($\beta_3 < 0$).	Negative association between coordination needs and delegation is mitigated in regions with higher social capital (i.e., lower self-centeredness of workers; $\beta_3 > 0$). In particular, when

			social capital is very high (respectively, coordination need is very low), the degree of delegation is always high regardless of the level of coordination needs (social capital).
Implications for development patterns	The above sorting pattern (“negative sorting”) promotes the development of industries with shorter production chains, but impedes the growth of industries with longer production chains, which have larger production and employment spillover.	Unequal skill distribution promotes the development of industries with shorter production chains, but impedes the growth of industries with longer production chains, which have larger production and employment spillover.	Delegation is positively associated with job satisfaction, wages (higher productivity), and skill upgrading of the workers. Thus, regions with higher social capital (lower self-centeredness of workers) may have a comparative advantage in industries requiring greater coordination and exhibit lower inequalities in terms of both income and skill.

(This page is intentionally left blank)

Chapter 2

Skill Sorting and Production Chains: Evidence from India*†

Abstract

Contrary to the theoretical prediction of most studies, this paper empirically shows that high-skilled workers are sorted into industries with shorter production chains in India. I hypothesize that such a reverse sorting pattern, which I call "negative sorting," occurs, because the returns to skill become lower in industries with longer production chains as a result of substantial quality deterioration. Such substantial quality deterioration along the production chains is likely to occur in developing countries such as India, which is characterized by a large pool of low-skilled labor, poor infrastructure, and less-advanced technology. Using both individual and industry-level data from the National Sample Surveys and India's input-output tables of 1999 and 2009, I find consistent evidence in favor of this hypothesized mechanism: High-skilled individuals are sorted into industries with shorter production chains, where returns to skill are higher. In addition, returns to skill are higher when the quality of intermediate inputs is better. The results remain robust even when correcting for possible selection bias and controlling for alternative reasons. The proposed mechanism provides one answer to the service-led growth of India, and makes a general contribution to the studies on inter-sector skill allocation.

JEL Classification: I25, I26, J24, J31, L23, O15

Keywords: India, Input quality, Production chains, Return to skill, Skill sorting

---

* Previous version of this chapter was released as the IDE discussion paper No. 545 (http://www.ide.go.jp/English/Publish/Download/Dp/545.html).

† I thank Yuji Genda, Hideaki Goto, Mizuki Harizuka, Masami Ishida, Satoshi Inomata, Md. Humayun Kabir, Ayako Kondo, Hisaki Kono, Tomohiro Machikita, Hiroaki Miyamoto, Mayumi Murayama, Mari Nakamura, Hideo Owan, Hiroshi Sasaki, Jeffrey M. Wooldridge, and all the internal seminar participants at the Institute of Developing Economies, Japan External Trade Organization (IDE-JETRO) and the University of Tokyo. I also greatly appreciate the helpful comments from Ryuichi Tanaka, a discussant for the previous version of this paper at the 2016 Spring Meeting of the Japanese Economic Association. This paper is the research output of the IDE-JETRO's FY2014–15 research project entitled "Transition from School to Work: Linkage of Education and Employment in India." I am grateful for the funding support from the IDE-JETRO.

The main text of Chapter 2 will be disclosed within five years.

For the earlier version of Chapter 2, see the IDE discussion paper No. 545 (http://www.ide.go.jp/English/Publish/Download/Dp/545.html).

Chapter 3

Skill Distribution and Comparative Advantage*†

Abstract

This chapter empirically examines different comparative advantages across countries in relation to their different skill distribution patterns. By utilizing industry export and skill distribution data on 58 economies around the world in 2000, I find that a country with a greater dispersion of skills exports relatively more in industries with shorter production chains, while a country with a more equal dispersion of skills has relatively higher exports in industries with longer production chains. The causal relationship is fairly robust against selection and endogeneity biases, and controlling for a country's average skill, infrastructure, and contract enforcement institutions.

JEL Classification: F14, F16, I25

Keywords: Comparative advantage, Production chains, Skill distribution

.

---

* This chapter is a substantial revision of Asuyama, Yoko (2012) "Skill Distribution and Comparative Advantage: A Comparison of China and India." *World Development*, 40(5): 956-969 (DOI: 10.1016/j.worlddev.2011.11.017).

† I am grateful to Yuji Genda, Hideaki Goto, Satoshi Inomata, Moriki Ohara, and three anonymous referees in *World Development* for helpful comments. I also thank IDE-JETRO for its financial support. All remaining errors are my own.

1. Introduction

What determines a country's comparative advantage? This paper focuses on one factor—a country's skill distribution—as a determinant of comparative advantage and aims to empirically identify the causal relationship. This paper is a substantial revision of Asuyama (2012), which is primarily motivated by the contrasting economic development paths of the two economic giants, China and India. Both China and India started economic reform around the early 1980s, when the income levels of these countries were almost the same.[1] However, their development patterns since then seem fairly different: China's growth has been driven by manufacturing, whereas India's has been fueled by services such as software, business process outsourcing, and call center services. In Asuyama (2012), I show that the different skill distribution patterns in these two countries (i.e., more equal skill distribution in China and unequal skill distribution in India) is one important factor that explains their different comparative advantages.[2]

Although the motivation, main idea, arguments, and basic empirical strategy are the same, this paper aims to substantially extend the analysis in Asuyama (2012) in the following ways: First, the main focus of this paper is not limited to China and India, but instead examines 58 economies around the world.[3] This increases the variation in countries' skill distribution and makes the empirical results more robust. Second, I directly link the hypothesized mechanism of this paper to the one introduced in Chapter

[1] Gross domestic product (GDP) per capita (current USD prices) in 1980 was 193 USD for China and 271 USD for India (World Bank 2014).

[2] For more details on the different comparative advantages and skill distribution patterns of China and India, see Asuyama (2012). Here, I just emphasize the stark contrast in their skill distribution patterns. As of 2010, the shares of employed people who were illiterate (Level 1), had completed less than primary, primary, and lower secondary education (Level 2), and upper secondary and post-secondary education (Level 3) were 32%, 42%, and 27% in India and 3%, 73%, and 24% in China. Furthermore, the proportion of the most high-skilled workers who have attained postgraduate (or above) education was 1.9% in India compared with 0.4% in China (COSC and NBS [2012] and NSSO [2011]).

[3] To strengthen the main results based on export data from China and India for 4 periods, Asuyama (2012) also examines export data from 103 countries. However, the analysis is very primitive and thus requires more sophisticated investigations.

2. Third, I employ a more standard estimation technique utilized in the studies exploring the sources of comparative advantage. Fourth, I provide more careful treatments with selection and endogeneity biases. Finally, I additionally control for a country's average skill level and contract enforcement institutions.

By utilizing industry export and skill distribution data on 58 economies in 2000, this paper empirically shows that a country with a relatively unequal skill distribution has relatively more exports in industry with shorter production chains, whereas a country with a relatively even skill distribution, has more exports in industries with longer production chains. This finding is fairly robust across different specifications, including those which correct for selection and endogeneity biases, and control for a country's average skill, infrastructure, and contract enforcement institutions.

In this paper, the length of production chains which varies across industry is defined by the amount of domestic intermediate inputs required in order to produce one dollar worth of industry output. It measures the size of production linkage with the domestic intermediate input industries. For example, automobile industry has longer production chains since producing a car requires a huge amount of parts, including both software and hardware. By contrast, software industry has shorter production chains, because writing a software program mainly requires labor and computers. As can be easily imagined from this simple example, production chains generally tend to be longer in manufacturing compared with agricultural, mining, and service sectors, although variation also exists within each sector (see Sections 4 and 7). Although the empirical results of this paper are based on nonservice industries owing to data unavailability, they imply that the difference in skill distribution between China and India has influenced the above-mentioned patterns of their comparative advantages.

How does the country's unequal skill distribution lead to a comparative advantage in industries with shorter production chains? The main hypothesis in this paper is based on the skill-sorting mechanism proposed in Chapter 2: C*eteris paribus*, in a country with greater skill dispersion, (i) a greater ratio of very low-skilled workers results in lower quality of each intermediate input; (ii) it then leads to greater quality deterioration in industries that require a larger amount of intermediate inputs (i.e., industries with longer production chains), as predicted by Kremer's (1993) O-ring theory; (iii) if such quality deterioration along the production chains is substantial, high-skilled workers are sorted into industries with shorter production chains, where their wages are less negatively affected by the lower-quality intermediate inputs ("negative sorting"). (iv) The probability of such a negative sorting and the relative skill gap between industries increases as the country's skill distribution becomes more unequal. (v) Assuming that industry productivity is determined by the skill intensity of the industry, an application of the Ricardian model to (iv) leads to the prediction that a country with a greater dispersion of skills exports more in industries with shorter production chains.

This study primarily contributes to the literature by proposing a new mechanism which links country skill distribution to comparative advantage via the length of industry production chains, and empirically identifying its causal relationship. This contributes to the studies exploring the sources of comparative advantage. As Chor (2010) summarizes the recent surge of empirical studies on sources of comparative advantages, differences between countries in productivity (as predicted by the Ricardian model), factor endowments (as predicted by the Heckscher-Ohlin model), and institutions have been identified as sources of comparative advantage (for more details, see the introduction of Chor (2010) and the papers cited).

With regard to the impact of skill distribution on comparative advantage, a few theoretical and empirical studies exist (Grossman and Maggi 2000; Grossman 2004; Ohnsorge and Trefler 2007; Bombardini *et al*., 2012). Although the idea of the present paper is closely related with Grossman (2004) and Bombardini *et al*. (2012) as explained in Section 2, none of the existing studies characterize industry by the length of production chains. In addition, among those studies, only Bombardini *et al.* (2012) provides empirical analysis. They theoretically and empirically show that a country with greater skill dispersion has more exports in industries with a lower degree of complementarity among worker's skill.

It turns out that my measure for the industry's length of production chains is positively correlated with the skill complementarity measure (or "skill substitutability" measure in their term) in their paper, although the correlation is weak.[4] The results of this study may thus partly capture the mechanism proposed by Bombardini *et al.* (2012). However, this study still has some advantages over their paper: First, my empirical analysis covers 58 economies, including both developed and developing ones, while Bombardini *et al.* (2012) cover only 19 developed countries. Second, like many other studies (e.g., Romalis 2004; Nunn 2007; Levchenko 2007; Tang 2012), Bombardini *et al.* (2012) construct the skill complementarity measure from the United States' (US) data and then apply it to all other countries. By contrast, the industry-specific production chain measure used in this paper is primarily constructed from each country's data. This treatment is more appropriate, given the technological differences across countries, in particular when analyzing both developed and developing countries. As a robustness check, I also apply a US-based production chain measure to all countries. Finally, this study provides more careful treatments with selection and endogeneity biases. I utilize two selection-correction methods that were proved to

[4] The correlation coefficient is 0.242 and insignificant at the 5% level (the correlation is calculated at the industry level for the 18 industries analyzed in the current paper). The skill complementarity index is *IndCoord2* in Chapter 4, which is the ranking of industries based on residual wage dispersion (measured by standard deviation) computed following Bombardini *et al.*'s (2012) method using the American Community Survey 2008–2012 5-year sample. In Chapter 4, my measure for the industry's length of production chains (*IndCoord1*) is more strongly correlated with *IndCoord2* (the correlation coefficient is 0.548 and significant at the 1% level). Besides the differences in years, this may reflect the following differences between the production chain measure (*ChainL*) in this chapter and that in Chapter 4: This chapter's *ChainL* does not cover service industries; it covers developing economies as well as developed economies; and it excludes the production chains for imported inputs.

perform very well in the Monte Carlo simulation.[5] I also explicitly correct for the endogeneity bias by instrumenting country skill dispersion.[6]

The rest of the paper is organized as follows. Section 2 presents hypothesized mechanisms. Section 3 explains the empirical methodology. Section 4 describes the data and explains the construction of the key variables. Sections 5 and 6 present the baseline estimation results and several analyses of robustness, respectively. Section 7 concludes.

2. Hypothesized Mechanisms

This section briefly presents the hypothesized mechanisms, which can explain why a country with a higher (respectively, lower) dispersion of skills has a comparative advantage in industries with shorter (longer) production chains.

The classical Ricardian model predicts that countries export relatively more in industries that are relatively more productive.[7] For example, consider a world with two countries (A and B), two industries (X and Y), and one production factor, labor. Country A exports relatively more in industry X if the equation below holds:

$$\frac{Z_{AX}}{Z_{BX}} > \frac{Z_{AY}}{Z_{BY}}, \quad \text{or equivalently,} \quad \frac{Z_{AX}}{Z_{AY}} > \frac{Z_{BX}}{Z_{BY}},$$

where Z_{ij} indicates labor productivity of country *i* in industry *j*. This Ricardian prediction is empirically confirmed by Costinot *et al*. (2012), who also developed a model with multiple countries, multiple industries, and one production factor (i.e.,

---

5 Bombardini *et al*. (2012) follow Helpman *et al*. (2008) and take a Heckman-based approach to correct for selection bias. But as Head and Mayer (2014: 178–179) mention, this approach only identifies intensive margin effects, while the two methods used in this paper identify both extensive and intensive margin effects combined. Estimating the total effects, including both extensive and intensive margins, seems more appropriate when examining a country's comparative advantage.

6 Bombardini *et al*. (2012) do not instrument a country's skill dispersion. They argue that their selection-correction method and the use of US data to construct a skill complementarity measure prevent the reverse causality. However, as mentioned in footnote 5, their selection-correction method has a drawback.

7 As for the Ricardian model, see textbooks on international economics, such as Feenstra (2004).

labor).

In a model with one production factor, which is labor, industry productivity is totally determined by labor productivity. Assuming that labor productivity is positively associated with workers' skill, the following prediction can be obtained: *Ceteris paribus*, countries export relatively more in industries that are relatively more skilled-labor-intensive.

How is the degree of skilled-labor intensity of industry (SK_{ij}) determined? The model in Chapter 2 provides one possible mechanism. In Chapter 2, I present a simple model in which industry is distinguished by the amount of necessary homogeneous intermediate inputs (i.e., "length of production chains"), and individuals who differ in their skill levels choose industries that offer the highest wages for their skills. When the quality of intermediate inputs degrades substantially as the length of production chains increases, high-skilled individuals choose industries with shorter production chains. Chapter 2 calls this sorting pattern "negative sorting" and the reverse pattern "positive sorting," which occurs when the quality deterioration is not sufficiently large (for more detail, see Chapter 2). When negative sorting (respectively, positive sorting) occurs, $SK_{iX} > SK_{iY}$ ($SK_{iX} < SK_{iY}$) holds, where industry X has shorter production chains than Y does.

In this paper, I explicitly assume that the degree of quality deterioration along the production chains (i.e., quality deterioration when larger amounts of intermediate inputs are combined together) is larger when the quality of each intermediate input (q in the model in Chapter 2) is lower. This assumption is reasonable: it also holds in Kremer's (1993) well-known O-ring production function.[8] In Chapter 2, q is

[8] This assumption can be expressed as $Q_{nq} > 0$ in the model in Chapter 2, where n is the amount of necessary intermediate inputs (length of production chains) and Q is the aggregated quality when n intermediates inputs with quality q are combined. In the O-ring production function (Kremer 1993: 553), the quality deterioration is captured by the term $\prod_{i=1}^{n} q_i$, where

exogenously determined by various factors, including levels of human capital, technology, and infrastructure of the economy. In this paper, I assume that *ceteris paribus*, q becomes lower as the ratio of "very low-skilled workers" is higher, where the "very low-skilled" are those with skill level below θ_L, which is lower than the average skill level of any countries.

With these assumptions, consider again the two-country, two-industry model in the beginning of this section. Industry Y has longer production chains (needs a greater amount of intermediate inputs to produce one unit of output) than X does. The productivity of each industry, Z_{ij} is determined by the degree of skilled-labor intensity in each industry. Each country has a uniform distribution of workers' skill with the same mean skill level $\bar{\theta}$, but country A has a more unequal skill distribution, that is, a mean-preserving spread of country B's skill distribution. Since the ratio of "very low-skilled workers" is larger in country A, q becomes lower, and thus the degree of quality deterioration along the production chains becomes larger in country A.

First, this result implies that negative sorting is more likely to occur in country A than in B. If country A exhibits negative sorting and country B exhibits positive sorting, country A exports more in industries with shorter production chains (industry X) because ($Z_{AX}/Z_{AY} > Z_{BX}/Z_{BY}$) holds as a result of $SK_{AX} > SK_{AY}$ and $SK_{BX} < SK_{BY}$. Second, if both countries exhibit negative sorting, both $SK_{AX} > SK_{AY}$ and $SK_{BX} > SK_{BY}$ hold. However, $SK_{AX}/SK_{AY} > SK_{BX}/SK_{BY}$ holds because the skill distribution of country A is more unequal. Then, country A has comparative advantage in industry X. In sum, it is predicted that a country with a higher (respectively, lower) dispersion of skills exports relatively more in industries with shorter (longer) production

---

q_i is skill (probability of perfect completion of task) of worker i and n is the number of tasks. Assuming $q_i = q \; \forall \, i$ and regarding $\prod_{i=1}^{n} q_i$ as Q, $Q_{nq} = (\partial^2 \prod_{i=1}^{n} q_i)/(\partial n \partial q) = nq^{n-1} \log n \geq 0$ holds.

chains as a result of skill sorting across industries.[9]

Alternative mechanisms are also possible. First, an industry's production chain length may proxy the industry-specific degree of imperfection of labor contracts in Grossman's (2004) model. In his two-country, two-industry model, high-skilled workers can receive wages according to their own skill levels in one industry (the "software" industry). However, their wages are dragged down by low-skilled workers in the other team-production industry ("automobile" industry) because of a imperfect labor contract. Then, high-skilled workers are sorted into the "software" industry. They also show that a country with greater skill dispersion exports more in the "software" industry.

Second, as mentioned in the Introduction section, this paper's measure on the length of production chains may partly represent the degree of industry-specific skill complementarity examined in Bombardini *et al*. (2012). They theoretically and empirically show that countries with higher skill dispersion export more in industries that exhibit lower skill complementarity. Different from the current paper and Grossman (2004), they assume that all industries inherit the skill distribution of the entire economy. Thus, countries with unequal skill distribution exhibit higher productivity in industries with higher skill substitutability (i.e., lower skill complementarity). Given the evidence on skill sorting across industries (Chapter 2; Asuyama and Goto 2016), however, their assumption that skill sorting does not occur seems too strong.

3. Empirical Strategy

I mostly follow the estimation strategy of Bombardini *et al*. (2012). Their estimation equation modifies the gravity equation, which aims to explain the size of

[9] The case in which both countries exhibit positive sorting is excluded because the empirical evidence in Chapter 2 and Asuyama and Goto (2016) show that at least some developing countries, such as India, exhibit negative sorting.

bilateral trade flows by various trade barriers as examined in Helpman *et al.* (2008). Similar empirical strategies are employed in several recent studies such as Cuñat and Melitz (2007), Levchenko (2007), Nunn (2007), and Chor (2010) which try to detect the sources of comparative advantage by estimating industry trade flows. In the baseline specification in this paper, I estimate the following equation:

$$\ln(Export_{xmi}) = \beta_1 ChainL_{xi} + \beta_2 ChainL_{xi} * SkillDisp_x + \gamma X_{xmi} + FE + \varepsilon_{xmi},$$

where $Export_{xmi}$ denotes the export value from exporter *x* to importer *m* in industry *i*. $ChainL_{xi}$ stands for the length of domestic production chain of industry *i* in exporter *x*. $SkillDisp_x$ measures the degree of skill dispersion of exporter *x*. X_{xmi} denotes a constant and other control variables as will be mentioned in Section 4.4. *FE* indicates various fixed effects: In most cases, they are exporter fixed effects, importer fixed effects, and industry fixed effects (FE1), but I also experiment with a combination of exporter fixed effects, and importer-industry fixed effects (FE2).

Our main focus is on the coefficient $\hat{\beta}_2$. As in other studies on comparative advantage (e.g., Nunn 2007; Bombardini et al. 2012), a negative $\hat{\beta}_2$ indicates that a country with a greater (respectively, lower) dispersion of skills exports relatively more in industries with shorter (respectively, longer) production chains, thus supporting my hypothesis. This can be expressed as $Export_{AX}/Export_{AY} > Export_{BX}/Export_{BY}$,[10] where exporter A has a more dispersed skill distribution than exporter B does, and the length of domestic production chains is shorter in industry X than in industry Y. (importer subscript *m* is omitted).[11]

[10] Equivalently, this condition can be expressed as either $Export_{BY}/Export_{BX} > Export_{AY}/Export_{AX}$, or $Export_{AX}/Export_{BX} > Export_{AY}/Export_{BY}$.

[11] As in other studies, a negative $\hat{\beta}_2$ does not necessarily indicate the following much stronger prediction: The *absolute* value of exports is higher in industries with shorter (respectively, longer) production chains in a country with a higher (lower) dispersion of skills. If $\hat{\beta}_1 > 0$ and $\hat{\beta}_2 < 0$ (and $\hat{\beta}_1 < |\hat{\beta}_2|$ in cases of using *SkillDispGini* and *SkillDispNonMid*, both of which range from 0 to 1), this prediction is also confirmed. As will become apparent in Tables 5.1 and

4. Data

4.1 Exports

The industry export flow data are from the "National Bureau of Economic Research-United Nations (NEBR-UN) Trade Data, 1962-2000" constructed by Feenstra and Lipsey (Feenstra *et al*. 2005). The sample is restricted to 58 exporters, which have information on skill, production chain length, and other control variables. The analysis focuses on year 2000, the same year as Bombardini *et al*. (2012). The original 4-digit Standard International Trade Classification (SITC Rev.2) codes in the export data are converted to the 18 industry classifications used for the data on production chain length.[12] Service exports are not covered in the regression analysis owing to data unavailability. The number of importers used in the present empirical analysis amounts to 161.

4.2 Exporter Skill Dispersion

Skill dispersion measure is constructed from the international educational attainment data constructed by Barro and Lee (Barro and Lee 2013). Using their distribution data on educational attainment for the populations over age 15 of each exporter, I construct three skill dispersion indices. The first is $SkillDispCV_x$, which is the coefficient of variation of skill computed as follows:

$$SkillDispCV_x = \sqrt{\sum_e [(YEDU_{ex} - SkillAvg_x)^2 * P_{ex}]} / SkillAvg_x ,$$

6.1, this stronger prediction is also confirmed in most cases. Since the past literature constructs industry characteristics from only one country, the effect of industry characteristics on exports ($\hat{\beta}_1$ in this paper) is absorbed by industry dummies and unidentified. Using country-industry-specific *ChainL* measure has an advantage in this aspect.

[12] Matching SITC Rev.2 with the 18 industry classification (that is based on ISIC Rev.3) is based on Arip *et al*. (2010) and World Bank's World Integrated Trade Solution (WITS) website (http://wits.worldbank.org/product_concordance.html)

where $SkillAvg_x$ is average years of education of the population in exporter x. Subscript e denotes the level of educational attainment (no schooling, primary, secondary, and tertiary). $YEDU_{ex}$ is the estimated average years of education for each schooling level.[13] P_{ex} denotes the ratio of population with educational level e. The second skill dispersion index is the Gini coefficient ($SkillDispGini_x$).[14] The third one is $SkillDispNonMid_x$, which is one minus the ratio of semi-skilled populations who have completed at least primary education but have not received any tertiary education. In order to minimize the possibility of reverse causality, following the treatment of Bombardini *et al*. (2012), I use the measure of skill dispersion in 1995, five years before the year in which the exports occur.

Appendix Table 1 arranges 58 exporters in ascending order of each $SkillDisp_x$ index and in descending order of $SkillAvg_x$. Although ranks of some countries (e.g., United States) substantially differ depending on which $SkillDisp_x$ index to use, these three indices are highly correlated in general (Table 4.1). $SkillDisp_x$ and $SkillAvg_x$ are strongly negatively correlated.

4.3 Industry Length of Production Chain

As an index for the length of production chains of industry i of exporter x ($ChainL_{xi}$), the column sum of the Leontief inverse coefficient of each industry computed from the year 2000's input-output (IO) tables of each country is used. The

---

[13] $YEDU_{ex} = 0$ for no schooling. For other educational level e, $YEDU_{ex}$ is exporter-specific average years of level-e education ($YEDUC_{ex}$) plus typical years of schooling to complete the previous educational stage ($YEDUP_e$, which are 6 years for completing primary and 12 years for secondary). $YEDUC_{ex}$ (which is reported in Barro and Lee [2013]) differs from the standard years required to complete level-e education, because the sample of $YEDUC_{ex}$ includes dropouts.

[14] Gini coefficient is computed by constructing a Lorenz curve, which plots the cumulative share of schooling (years) attained and the cumulative share of people from lowest to highest education levels.

data are taken from the OECD Input-Output Database (2015 edition) (OECD 2015).[15] This $ChainL_{xi}$ measures the amount of domestic intermediate input industry *i* requires, both direct and indirect, to produce one dollar's worth of output in industry *i*.[16] I use this $ChainL_{xi}$ as a proxy for the length of production chains of industry *i*. It should be noted that only domestic inputs are used to compute $ChainL_{xi}$, since the quality of imported input is assumed to be relatively good and is not affected by the skill of domestic workers. Industry's intensity of imported input usage is separately controlled for in regression analyses.

Table 4.2 displays $ChainL_{xi}$ by 18 industries. It first reports the average $ChainL_{xi}$ for the 58 exporters (weighted by export value). It also reports the rank of $ChainL_{xi}$ for several relatively large exporters, which are characterized by lower skill dispersion (Netherlands, Sweden, and Japan) and higher skill dispersion (India, Indonesia, and Brazil). US and China are placed in the middle because the relative position of these two countries depend on which skill index is used. There is substantial variation in the rank of $ChainL_{xi}$ across exporters, indicating the adequacy of using country-specific industry characteristics measure. Such variation is likely to be because

[15] The OECD database provides the actual figures for $ChainL_{xi}$, but $ChainL_{xi}$ can be calculated as follows: $ChainL_{xi} = \sum_k leon_{xki}$, where $leon_{xki}$ is the Leontief inverse coefficient in cell *ki* of exporter *x*'s IO table, and subscripts *k* and *i* denote row and column of the IO table, respectively. The Leontief inverse coefficient matrix *L* comprised of *k***i* $leon_{xki}$ s is computes as $L = (I - A_d)^{-1}$, where I is the identity matrix; A_d is the input coefficient matrix for domestic inputs, in which the coefficient in cell *ki* is the amount of domestic input of industry *k* used by industry *i* divided by the output of industry *i*.

[16] For example, suppose that in order to produce one unit of output, an automobile industry directly uses 0.4 units of input from the automobile industry itself, 0.2 units from the steel industry, and 0.1 units from the computer industry (the remaining 0.3 units are value added). Consequently, the 0.4 units of input from the automobile industry further require 0.4*0.4 units of input from the automobile industry, 0.4*0.2 units from steel industry, and 0.4*0.1 units from computer industry. Again, to produce the 0.4*0.4 units of input from the automobile industry, the automobile industry requires 0.4*0.4*0.4 inputs from the automobile industry itself….and so on. In this way, one output generated by an industry also indirectly generates chains of demand for intermediate inputs.

of the differences in industry technology and the degree of dependence on imported inputs. It is possible that firms in countries with higher skill dispersion have more incentives to shorten the length of domestic production chains in order to reduce the negative impact from low-skilled workers by increasing imported inputs. In this case, however, the impact of country's skill distribution on comparative advantage is further strengthened through affecting firms' decision on the length of domestic production chains.

4.4 Other Control Variables

Appendix Table 2 provides the detailed explanations, data sources, and summary statistics for the additional control variables. The first group of them accounts for various conventional trade barriers between exporter–importer pairs. They include logarithm of distance ($Distance_{xm}$); the presence of colonial ties ($Colonial_{xm}$); geographically contiguity ($Contiguous_{xm}$); shared legal systems, languages, and religions ($Legalsystem_{xm}$, $Language_{xm}$, and $Religion_{xm}$); the number of exporter/importer who are members of the World Trade Organization (WTO_{xm}); shared regional trade arrangement (RTA_{xm}) or a currency (CU_{xm}); and the number of islands and landlocked countries in exporter-importer pair ($Islands_{xm}$, and $Landlocked_{xm}$).

The second group of control variables includes endowment characteristics of the exporting country and its industries. These control variables include ratio of imported inputs of industry ($Importr_{xi}$); interaction term of capital intensity of exporter (K_x) with that of industry (K_i); interaction term of skill intensity of exporter (SK_x = $SkillAvg_x$) with that of industry (SK_i). These two interaction terms (which are only available for manufacturing industries) are added to control for effects predicted by the Heckscher-Ohlin model that a country exports relatively more in industries using relatively more abundant factors in the country. Note that K_i and SK_i are computed from US data, and thus common across all countries as in other studies (e.g. Chor 2010;

Bombardini *et al*. 2012).

5. Baseline Estimation Results

Table 5.1 reports the baseline OLS regression results with the logarithm of bilateral industry exports as the dependent variable. The estimates in the first four columns use all exports in 18 industries. However, exports from primary (agriculture, forestry, fishing, and mining) industries might be affected by inputs such as land, weather, and natural resources, which are not included as inputs in the IO tables nor affected by worker skill levels. Considering those unobserved factors on the primary industries, the remaining four columns restrict the sample to only the 16 manufacturing industries.

Consistent with my hypothesis, the estimated coefficients for *ChainL*SkillDisp* ($\hat{\beta}_2$) are negative and statistically significant in all specifications in all-industry sample, regardless of skill dispersion indices and the types of fixed effects. As for the manufacturing sample, however, the results are not so robust. Although $\hat{\beta}_2$ tends to be negative when *SkillDispCV* or *SkillDispGini* is used, most of the estimates are statistically insignificant when *ChainL*SkillAvg* is controlled for. When *SkillDispNonMid* is used and *ChainL*SkillAvg* is controlled for, $\hat{\beta}_2$ is even significantly positive.

The results for other variables are similar regardless of the skill dispersion indices, although the results obtained when using *SkillDispGini* or *SkillDispNonMid* are not displayed. The estimated coefficients for standard trade barriers between exporter-importer pairs mostly exhibit the predicted signs. Distance and the number of landlocked countries are negatively associated with exports. Geographic continuity, colonial ties, common legal systems, languages, and religions, and regional trade agreement (RTA) are all positively associated with exports. The estimated coefficients on WTO membership and common currency tend to be positive, although not all estimates are statistically significant. Somewhat unexpectedly, industry ratio of imported input is negatively associated with exports. This might be because greater dependence on imported inputs captures the lower competitiveness of the industry. The estimated positive relationships between exports and the interaction terms of capital/skill intensity of exporting country and industry confirm the prediction of the Heckscher-Ohlin model. Finally, consistent with the predictions by Kremer (1993) and Costinot *et al*. (2013), the coefficient of *ChainL*SkillAvg* tends to be positive: that is, countries with higher average skill of workers export relatively more in industries with longer production chains.

6. Robustness Analysis

6.1 Selection Corrections

The baseline estimations omit observations with export values of zero, due to their log transformation. However, observations with zero exports constitute 61.7% of the total exporter-importer-industry cells in the sample. Zero bilateral industry trade is due to either data recording practices (e.g., rounding or declaration thresholds) or the presence of very high trade costs (Head and Mayer 2014: 177-178). Excluding zero export observations may generate biased estimates by introducing measurement errors or correlation between observed and unobserved trade barriers (Helpman *et al*. 2008). To correct for such bias, I utilize two methods, both of which perform very well in the Monte Carlo simulation by Head and Mayer (2014: 177-183).[17] They are the Poisson pseudo-maximum-likelihood (PPML) method recommended by Silva and Tenreyro

[17] PPML performs best when the error term is CVMR ("a Poisson-like error with a constant variance to mean ratio"), while EK Tobit performs best when the error term is log-normal.

(2006),[18] and the EK Tobit method proposed by Eaton and Kortum (2001).[19]

Table 6.1 reports the PPML and EK Tobit estimation results.[20] The estimated coefficients for *ChainL*SkillDisp* ($\hat{\beta}_2$) is always significantly negative in all-industry sample. $\hat{\beta}_2$ is also significantly negative in manufacturing sample in all specifications but the following two cases: (i) PPML estimators using *SkillDispGini*, and (ii) EK Tobit estimator using *SkillDispNonMid*. In these cases, $\hat{\beta}_2$ is ether insignificant or even positive. The signs of the estimated coefficients for the other control variables are almost the same as those in Table 5.1. The exception is that common currency dummy is now negatively associated with exports in all specifications. In addition, the coefficient of *ChainL*SkillAvg* is insignificant or even negative when using *SkillDispCV*, although it still tends to be positive when using other skill dispersion measures.

6.2 Endogeneity

The patterns of industry exports may influence the skill distribution and the length of industry production chains of the exporting country. In addition to such a reverse causality problem, both $SkillDisp_x$ and $ChainL_{xi}$ might be correlated with unobserved factors in the error term. Such endogeneity problems would lead to the

[18] I use "ppml" command of Stata developed by Silva and Tenreyro. The PPML estimator is similar to the ordinary Poisson estimator but has advantages over it in terms of dealing with convergence and numerical problems (Silva and Tenreyro 2011).

[19] Following the Stata code by Head and Mayer (see the website below), EK Tobit is estimated using "intreg" command in Stata, which is used for interval regression (generalized model of Tobit). The first dependent variable is log of export where observations with zero exports are missing as in baseline regression, while the second dependent variable is log of export ($Export_{xmi}$) where zero exports are replaced with the importer-industry-specific minimum positive export $\underline{Export}_{mi}$. This $\underline{Export}_{mi}$, which is the importer-industry-specific censoring point, is the hypothesized threshold below which disaggregated data by exporters are not reported (Eaton and Kortum 2001).
(https://sites.google.com/site/hiegravity/stata-programs)

[20] Since the sign and size of estimates are similar regardless of incorporating FE1 or FE2 in Table 5.1, and controlling for FE2 is computationally hard particularly in PPML and EK Tobit estimations, FE1 is always controlled for in the subsequent tables.

biased estimates for the coefficients of $ChainL_{xi} * SkillDisp_x$ in Tables 5.1 and 6.1.

To deal with the endogeneity problem, $SkillDisp_x$, which represents the skill dispersion in 1995, is instrumented with the corresponding $SkillDisp_x$ in 1950.[21] As for $ChainL_{xi}$, I follow the approach by Bombardini *et al.* (2012) and replace country-industry-specific *ChainL* with US-industry-specific production chain measure, $ChainLUS_i$. This $ChainLUS_i$, which is extracted from OECD (2015), measures the length of production chain of industry *i*, computed based on total inputs (including both domestic and imported ones). Observations in which US is the exporter is excluded from the regression sample. Consequently, $ChainLUS_i$ is completely exogenous to the remaining 57 exporters.

Table 6.2 reports several endogeneity-corrected estimates.[22] The first two columns contain the two-stage least squares (2SLS) estimates, where the dependent variable is the logarithm of exports as in Table 5.1, excluding the zero export observations. In columns (3) and (4), EK Tobit estimates are corrected by control function approach, following Wooldridge (2010: 784): In the first stage, each potentially endogenous regressor ($ChainLUS_{xi} * SkillDisp_x$ and $ChainLUS_{xi} * SkillAvg_x$) is regressed on all the excluded instruments and exogenous covariates. In the second stage, the obtained two residuals from the first-stage regressions ($\hat{v}_{1xmi}$ and $\hat{v}_{2xmi}$), which control for the endogeneity, are added to the explanatory variables of EK Tobit estimation in Table 6.1. If the coefficients of $\hat{v}_{1xmi}$, and $\hat{v}_{2xmi}$ are jointly statistically different from zero, $SkillDisp_x$, and $SkillAvg_x$ need to be treated as endogenous.

As the Kleibergen-Paap Wald rk F statistic ("Weak IV test stat.") in Table 6.2 suggests that our instruments are strongly correlated with the potentially endogenous

[21] $SkillAvg_x$ is also instrumented with $SkillAvg_x$ in 1950. The skill data in 1950 are taken from Barro and Lee (2013).
[22] Endogeneity-corrected Poisson estimates are not reported due to a lack of convergence when using "ivpoisson" command in Stata.

regressors in all specifications. Furthermore, the results on the endogeneity test indicate that $SkillDisp_x$, and $SkillAvg_x$ need to be treated as endogenous. Again, the estimated coefficients for *ChainLUS*SkillDisp* ($\hat{\beta}_2$) are always significantly negative in all-industry sample. $\hat{\beta}_2$ is also significantly negative in the manufacturing sample, except one insignificant result (2SLS estimator when using *SkillDispNonMid*). A positive $\hat{\beta}_2$ is no longer observed after correcting for endogeneity.

As for the other control variables, the estimated coefficient of *ChainL*SkillAvg* is now always significantly negative. This result is difficult to interpret and needs further scrutiny. Bombardini *et al*. (2012) also find negative effect of the average skill level of the economy on the exports in industries with lower skill complementarity. As mentioned in Introduction, my *ChainL* is weakly but positively correlated with their skill complementarity measure. It is interesting that both studies find similar effects of the average skill level of the economy.

Table 6.3 compares the magnitude of $\hat{\beta}_2$ across Tables 5.1, 6.1, and 6.2. The statistic in Table 6.3 displays [exp($\hat{\beta}_2$*SD)-1]*100(%), where *SD* is one standard deviation of *SkillDisp* (0.212 for *SkillDispCV*, 0.120 for *SkillDispGini*, and 0.105 for *SkillDispNonMid*, respectively). The magnitude of $\hat{\beta}_2$ tends to be much larger when correcting for selection and endogeneity biases. As an example, consider the most modest (and significantly negative) estimate of -15.2%. In this case (EK Tobit using *SkillDispNonMid* in all-industry sample), a unit increase in *ChainL* is associated with 32.4% (=[exp($\hat{\beta}_1$)-1]*100(%), where $\hat{\beta}_1$=0.281) increase in exports, although $\hat{\beta}_1$ is insignificant (Table 6.1). However, this positive effect would be 15.2% lower if country's skill distribution increases by one standard deviation. To put it differently, this implies that everything else equal, if the *SkillDispNonMid* (0.576) of India, a country with a very high skill dispersion, had been the same as that of China (0.328), its exports of "motor vehicles, trailers and semi-trailers (India's *ChainL* is 2.325)" would increase from 498 to 1228 million USD.[23]

6.3 Alternative Factors Controlled

One of the key hypotheses of this paper is that the productivity of an industry with longer production chains is more likely to be dragged down by the very low-skilled workers involved across the chains. However, negative impacts resulted from poor quality of production infrastructure, may also accumulate across the production chains. In addition, it is possible that the production chain measure used in this paper partly captures the product complexity measure in Levchenko (2007).[24] Then, it is necessary to control for the quality of a country's contract enforcement institutions, since Levchenko (2007) find that countries with better contracting institutions have a comparative advantage in sectors producing more complex products.

In order to control for these alternative factors, I added the interaction term between them ($AltFactor_x$) and $ChainLUS_{xi}$ to the regressions in Table 6.2. As a measure for the infrastructure quality, two indicators are used: One is $Eleloss_x$, which is the electric power transmission and distribution loss in the exporting country (measured as a ratio to output). The other is $Roadp_x$, which is the ratio of paved road to total roads. As for the quality of contract enforcement institutions, the rule of law index ($RoLaw_x$) developed by Kaufman *et al.* (2010) is used as in Levchenko (2007).

Table 6.4 reports the estimation results. The coefficient of *ChainLUS*SkillDsip* ($\hat{\beta}_2$) is always significantly negative regardless of estimation methods, industry coverage, skill dispersion indices, and the *AltFactor* variables. The coefficient of

---

[23] The calculation is as follows: First, the logarithm of predicted exports, 1,227,878 thousand USD is 14.021. This predicted log export is computed as ln(actual export, which is 497,732 USD)+ $\hat{\beta}_2$ **ChainL**(*SkillDispNonMid* of China – *SkillDispNonMid* of India)=13.118+(-1.566)*2.325*(0.328-0.576)=14.021.

[24] The focus of Levchenko (2007) is on incomplete contract. His industry-specific product complexity measure is thus constructed to measure the number of major suppliers. It is (-1)*[Herfindahl index of intermediate input use] of each industry.

*ChainLUS*SkillAvg* is again almost always negative. As expected, in all-industry sample, *ChainLUS*Eleloss* tend to be negatively associated with exports, while *ChainLUS*Roadp* and *ChainLUS*RoLaw* tend to be positively associated with exports. These associations between *ChainLUS*AltFactor* and exports mostly reverse in manufacturing sample, but this puzzling result is left for future research.[25]

7. Concluding Remarks

This paper empirically examines the country's different comparative advantages which result from the different skill distribution patterns in each country. By utilizing industry export data on 58 economies in 2000, this paper finds that a country with a greater dispersion of skills exports relatively more in industries with shorter production chains. Conversely, a country with a more equal dispersion of skills is found to have relatively higher exports in industries with longer production chains. The causal relationship is fairly robust across different specifications, including those which correct for selection and endogeneity biases, and control for a country's average skill, infrastructure, and contract enforcement institutions.

The main estimation results are only strictly applicable to nonservice industries. However, as Appendix Table 3 shows, the length of domestic production chains tends to be shorter in service industries than in manufacturing industries. Therefore, our results are consistent with the two-country comparison illustrated in Asuyama (2012): China, a country with relatively narrower dispersion of skills, has a comparative advantage in large-scale manufacturing industries with longer production chains, while India, a country with a greater dispersion of skills, has a comparative advantage in service

[25] When endogeneity is not corrected, the coefficient of *ChianLUS*RoLaw* is always significantly positive even in manufacturing sample, although the results do not change in the cases of *ChainLUS*Eleloss* and *ChainLUS*Roadp*.

industries with shorter production chains.

Production in industries with longer domestic production chains is beneficial to the economy by inducing larger demand for intermediate input production, which generally generates more employment. Thus, if developing countries, such as India, for example, would like to foster industries with long domestic production chains and increase exports in these industries, it needs to increase the number of semi-skilled workers with primary or secondary education and make skill distribution more equal. As Asuyama (2011) has examined, potential solutions may include various reforms in education and training policies, such as redesigning the financing system for education and the incentive structure for teachers and local government officials, as well as simply improving the quantity and quality of primary and secondary education.

Several areas remain for future research. First, the unexpected negative effects of the average skill, infrastructure, and contract enforcement institutions on the exports in industries with longer production chains, most of which become evident after correcting for endogeneity, need further scrutiny. Second, incorporating service industries into the empirical analysis is necessary. Finally, it is also important to empirically check whether the skill-sorting mechanism mentioned in Section 2 is really responsible for the generation of comparative advantage.

References

Arip, M. Affendy, Sim Yee Lau, and Satoru Madono. 2010. “Commodity-Industry Classificationproxy: A Correspondence Table between SITC Revision 2 and ISIC Revision 3.” MPRA Paper 27626, University Library of Munich, Germany.

Asuyama, Yoko. 2011. “Skill Formation through Education and Training: A Comparison of China and India.” In Moriki Ohara, M. Vijayabaskar, and Hong Lin eds., *Industrial Dynamics in China and India: Firms, Clusters, and Different Growth*

Paths. Basingstoke, Hampshire: Palgrave Macmillan, Chapter 5, 107-134.

Asuyama, Yoko. 2012. "Skill Distribution and Comparative Advantage: A Comparison of China and India." *World Development*, 40(5): 956-969.

Asuyama, Yoko, and Hideaki Goto. 2016. "Negative Skill Sorting Across Production Chains." IDE Discussion Paper No.557.

Barro, Robert J., and Jong-Wha Lee. 2013. "A New Data Set of Educational Attainment in the World, 1950-2010." *Journal of Development Economics*, 104:184-198.

Barro, Robert J., and Rachel M. McCleary. 2005. "Which Countries Have State Religions?" *Quarterly Journal of Economics*, 120(4), 1331–1370.

Baum, Christopher F., Mark E. Schaffer, and Steven Stillman. 2007. "Enhanced Routines for Instrumental Variables/Generalized Method of Moments Estimation and Testing." *STATA Journal*, 7(4): 465-506.

Becker, Randy, Wayne Gray, Jordan Marvakov. 2014 (Revised April 2016) NBER-CES Manufacturing Industry Database (1958-2011). Retrieved from http://www.nber.org/nberces/

Bombardini, Matilde, Giovanni Gallipoli, and Germán Pupato. 2012. "Skill Dispersion and Trade Flows." *American Economic Review*, 102(5): 2327–2348.

CEPII (Centre D'Etudes Prospectives et D'Informations Internationales). *Distances Measures*. Retrieved from http://www.cepii.fr/anglaisgraph/bdd/distances.htm

COSC and NBS (Census Office under the State Council, Department of Population and Employment Statistics, and National Bureau of Statistics) 2012. Tabulation on the 2010 population census of the People's Republic of China. Beijing: China Statistics Press.

Chor, Davin. 2010. "Unpacking Sources of Comparative Advantage: A Quantitative Approach." *Journal of International Economics*, 82(2), 152–167.

Costinot, Arnaud, Dave Donaldson, and Ivana Komunjer. 2012. "What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas." *Review of*

Economic Studies, 79: 581–608.

Costinot, Arnaud, Jonathan Vogel, and Su Wang. 2013. "An Elementary Theory of Global Supply Chains." *Review of Economic Studies*, 80(1): 109-144.

Cuñat, Alejandro, and Marc J. Melitz. 2007. “Volatility, Labor Market Flexibility, and the Pattern of Comparative Advantage.” NBER Working Paper No. 13062.

Eaton, Jonathan, and Samuel Kortum. 2001. “Trade in Capital Goods.” *European Economic Review*, 45(7): 1195-1235.

Feenstra, Robert C. 2004. *Advanced International Trade: Theory and Evidence*. Princeton and Woodstock: Princeton University Press.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. 2015. “The Next Generation of the Penn World Table.” *American Economic Review*, 105(10): 3150–3182.

Feenstra, Robert. C., Robert E. Lipsey, Haiyan Deng, Alyson C. Ma, and Hengyong Mo 2005. “World Trade Flows: 1962-2000.” NBER Working Paper, No. 11040.

Grossman, Gene M. 2004. “The Distribution of Talent and the Pattern and Consequences of International Trade.” *Journal of Political Economy*, 112(1): 209-239.

Grossman, Gene M., and Giovanni Maggi. 2000. “Diversity and Trade.” *American Economic Review*, 90(5): 1255-1275.

Head, Keith, and Thierry Mayer. 2014. “Gravity Equations: Workhorse, Toolkit, and Cookbook.” In Gita Gopinath, Elhanan Helpman and Kenneth Rogoff eds., *Handbook of International Economics*, Volume 4. Chapter 3, 131-195.

Head, Keith, Thierry Mayer, and John C. Ries. 2010. “The Erosion of Colonial Trade Linkages After Independence.” *Journal of International Economics*, 81(1), 1-14.

Helpman, Elhanan, Marc Melitz, and Yona Rubinstein. 2008. “Estimating Trade Flows: Trading Partners and Trading Volumes.” *Quarterly Journal of Economics*, 123(2): 441-487.

Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi. 2010. “The Worldwide

Governance Indicators: A Summary of Methodology, Data and Analytical Issues." World Bank Policy Research Working Paper No. 5430.

Kremer, Michael. 1993. "The O-Ring Theory of Economic Development." *Quarterly Journal of Economics*, 108(3): 551-575.

Levchenko, Andrei A. 2007. "Institutional Quality and International Trade." *Review of Economic Studies*, 74(3): 791–819.

Mayer, Thierry and Soledad Zignago. 2006. "Notes on CEPII's Distances Measures." Retrieved from http://www.cepii.fr/anglaisgraph/bdd/distances.htm

NSSO (National Sample Survey Office, National Sample Survey Office, Ministry of Statistics & Programme Implementation, Government of India). 2011. *Employment and Unemployment Situation in India 2009-10: NSS 66th Round*. New Delhi: NSSO.

Nunn, Nathan. 2007. "Relationship-Specificity, Incomplete Contracts, and the Pattern of Trade." *Quarterly Journal of Economics*, 122(2): 569-600.

Ohnsorge, Franziska, and Daniel Trefler. 2007. "Sorting It Out: International Trade with Heterogeneous Workers." *Journal of Political Economy*, 115(5): 868-892.

OECD (Organisation for Economic Co-operation and Development). 2015. Input-Output Database, 2015 edition.

Silva, Santos J.M.C., and Silvana Tenreyro. 2006. "The Log of Gravity." *Review of Economics and Statistics*, 88(4): 641–658.

Silva, Santos J.M.C., and Silvana Tenreyro. 2011. "Poisson: Some Convergence Issues." *STATA Journal*, 11(2): 207-212.

Tang, Heiwai. 2012. "Labor Market Institutions, Firm-Specific Skills, and Trade Patterns." *Journal of International Economics*, 87(2): 337–351.

Wooldridge, Jeffrey. M. 2010. *Econometric Analysis of Cross Section and Panel Data, Second Edition*. MIT Press.

World Bank. 2014. *World Development Indicators* (22 July, 2014 version).

Table 4.1 Correlations among Three *SkillDisp* Indices and *SkillAvg*

	SkillDispCV	*SkillDispGini*	*SkillDispNonMid*	*SkillAvg*
SkillDispCV	1			
SkillDispGini	0.927	1		
SkillDispNonMid	0.764	0.657	1	
SkillAvg	-0.934	-0.819	-0.669	1

Notes: Correlations are calculated at the country level. All correlations are statistically significant at 1 % level.

Table 4.2 Length of Production Chains (*ChainL*) by 18 Industries

Industry	58 exporters			Nether-lands	Swe-den	Japan	US	China	Brazil	Indo-nesia	India
	Mean	Std. Dev.	Rank	Rank							
Mining and quarrying	1.367	0.255	1	1	10	4	1	2	2	1	1
Coke, refined petroleum products and nuclear fuel	1.661	0.264	2	4	1	1	9	6	18	4	3
Agriculture, hunting, forestry and fishing	1.748	0.200	3	15	2	2	14	1	1	2	2
Computer, Electronic and optical equipment	1.773	0.205	4	10	14	10	4	3	8	14	10
Other non-metallic mineral products	1.829	0.247	5	11	12	3	6	7	6	5	4
Fabricated metal products	1.879	0.281	6	5	8	9	5	14	7	6	8
Other transport equipment	1.885	0.263	7	12	4	15	2	16	14	7	16
Pulp, paper, paper products, printing and publishing	1.891	0.169	8	16	15	8	11	11	4	12	6
Manufacturing nec; recycling	1.892	0.214	9	2	16	16	3	4	5	18	13
Chemicals and chemical products	1.897	0.243	10	17	3	13	13	10	12	8	11
Machinery and equipment, nec	1.900	0.237	11	6	7	12	7	12	10	3	7
Electrical machinery and apparatus, nec	1.923	0.330	12	7	9	11	8	15	13	9	9
Rubber and plastics products	1.957	0.275	13	13	6	14	12	13	16	10	17
Basic metals	1.966	0.304	14	3	11	17	15	18	11	11	5
Wood and products of wood and cork	1.992	0.182	15	14	18	5	17	9	3	15	12
Textiles, textile products, leather and footwear	2.044	0.323	16	9	5	6	10	8	9	16	18
Motor vehicles, trailers and semi-trailers	2.045	0.350	17	8	13	18	16	17	15	13	15
Food products, beverages and tobacco	2.134	0.162	18	18	17	7	18	5	17	17	14

Notes: Statistics for 58 exporters are weighted by export value. Rank is attached in ascending order of *ChainL*. Highlighted two industries are non-manufacturing industries.

Source: OECD (2015).

Table 5.1 Determinants of Comparative Advantage: Baseline OLS Results

	All				Manufacturing			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
SkillDisp = SkillDispCV								
ChainL	1.648***	1.574***	0.227	0.093	1.901***	1.873***	0.610	0.815
	(0.132)	(0.130)	(0.524)	(0.517)	(0.151)	(0.147)	(0.708)	(0.696)
*ChainL*SkillDisp*	-1.824***	-1.788***	-0.959***	-0.888***	-1.644***	-1.647***	-0.780	-0.938*
	(0.173)	(0.168)	(0.348)	(0.342)	(0.234)	(0.226)	(0.523)	(0.504)
*ChainL*SkillAvg*			0.111***	0.116***			0.097*	0.080
			(0.041)	(0.040)			(0.052)	(0.052)
Importr	-0.783***	-0.924***	-0.776***	-0.917***	-0.397***	-0.501***	-0.385***	-0.491***
	(0.134)	(0.131)	(0.134)	(0.131)	(0.148)	(0.143)	(0.149)	(0.144)
*Kx*Ki*					0.120***	0.139***	0.116***	0.135***
					(0.016)	(0.015)	(0.016)	(0.015)
*SKx*Ski*					0.454***	0.477***	0.465***	0.487***
					(0.037)	(0.036)	(0.038)	(0.037)
Distance	-0.965***	-0.992***	-0.965***	-0.993***	-1.003***	-1.025***	-1.003***	-1.025***
	(0.023)	(0.024)	(0.023)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
Contiguous	0.469***	0.469***	0.468***	0.469***	0.429***	0.433***	0.429***	0.432***
	(0.075)	(0.077)	(0.075)	(0.077)	(0.078)	(0.080)	(0.078)	(0.080)
Legalsystem	0.251***	0.265***	0.251***	0.265***	0.266***	0.279***	0.266***	0.279***
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.027)	(0.026)	(0.027)
Colonial	0.665***	0.673***	0.664***	0.673***	0.679***	0.682***	0.679***	0.682***
	(0.075)	(0.076)	(0.075)	(0.076)	(0.079)	(0.079)	(0.079)	(0.079)
Language	0.220***	0.233***	0.221***	0.233***	0.250***	0.261***	0.250***	0.261***
	(0.041)	(0.042)	(0.041)	(0.042)	(0.042)	(0.043)	(0.042)	(0.043)
Religion	0.395***	0.398***	0.395***	0.398***	0.443***	0.448***	0.443***	0.448***
	(0.058)	(0.059)	(0.058)	(0.059)	(0.059)	(0.061)	(0.059)	(0.061)
WTO	0.472	0.606	0.623	0.762	-0.820	0.805***	-4.962***	1.008***
	(0.601)	(0.586)	(0.599)	(0.584)	(0.693)	(0.273)	(0.160)	(0.300)
RTA	0.231***	0.221***	0.231***	0.221***	0.244***	0.237***	0.244***	0.237***
	(0.047)	(0.048)	(0.047)	(0.048)	(0.049)	(0.050)	(0.049)	(0.050)
CU	0.204**	0.137	0.203**	0.136	0.198*	0.134	0.198*	0.134
	(0.103)	(0.106)	(0.103)	(0.106)	(0.110)	(0.112)	(0.110)	(0.112)
Islands	-2.258***	2.256***	-2.526***	2.233***	-3.005***	-3.286***	-3.926***	2.011***
	(0.163)	(0.576)	(0.188)	(0.577)	(0.192)	(0.182)	(0.206)	(0.233)
Landlocked	-1.115***	-1.118***	-1.328***	-1.341***	-2.742***	-1.559***	-3.026***	-1.687***
	(0.115)	(0.116)	(0.139)	(0.140)	(0.163)	(0.136)	(0.226)	(0.162)
SkillDisp = SkillDispGini								
ChainL	1.719***	1.641***	0.053	-0.049	1.864***	1.828***	0.049	0.057
	(0.140)	(0.137)	(0.382)	(0.377)	(0.159)	(0.154)	(0.441)	(0.440)
*ChainL*SkillDisp*	-3.437***	-3.365***	-1.663***	-1.566***	-2.711***	-2.686***	-0.737	-0.762
	(0.342)	(0.332)	(0.514)	(0.504)	(0.428)	(0.412)	(0.601)	(0.586)
*ChainL*SkillAvg*			0.131***	0.133***			0.141***	0.137***
			(0.029)	(0.029)			(0.033)	(0.033)

SkillDisp = SkillDispNonMid								
ChainL	1.811***	1.725***	-0.027	-0.128	1.218***	1.138***	-1.401***	-1.514***
	(0.149)	(0.148)	(0.294)	(0.294)	(0.165)	(0.164)	(0.341)	(0.343)
*ChainL*SkillDisp*	-3.738***	-3.646***	-1.932***	-1.819***	-0.595	-0.435	1.717***	1.910***
	(0.394)	(0.388)	(0.466)	(0.464)	(0.462)	(0.457)	(0.529)	(0.528)
*ChainL*SkillAvg*			0.153***	0.153***			0.219***	0.222***
			(0.023)	(0.023)			(0.026)	(0.026)
Fixed effects	FE1	FE2	FE1	FE2	FE1	FE2	FE1	FE2
Observations	65858	65858	65858	65858	59946	59946	59946	59946
Adj. R-squared	0.586	0.614	0.586	0.614	0.616	0.644	0.616	0.644

Notes: The dependent variable is the logarithm of exports from exporter *x* to importer *m* in industry *i*. Robust standard errors, clustered by exporter-importer pair are reported in parentheses. FE1 indicates exporter fixed effects (FE), importer FE, and industry FE. FE2 indicates exporter FE, and importer-industry FE. Although not reported, in the regressions using *SkillDispGini* and *SkillDispNonMid*, the same control variables as those when using *SkillDispCV* are included. The adjusted R-squared is for the case using *SkillDispCV*, but the statistic is almost the same for the cases using *SkillDispGini* or *SkillDispNonMid*. ***p<0.01, **p<0.05, *p<0.1.

Table 6.1 Selection Corrected Estimates: PPML and EK Tobit Estimations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	PPML				EK Tobit			
	All		Manufacturing		All		Manufacturing	
SkillDisp=SkillDispCV								
ChainL	1.522***	1.957	2.153***	4.943*	2.512***	3.462***	2.535***	4.039***
	(0.481)	(1.976)	(0.640)	(2.791)	(0.132)	(0.544)	(0.156)	(0.755)
*ChainL*SkillDisp*	-2.655***	-2.929**	-3.012***	-5.008**	-2.206***	-2.776***	-1.441***	-2.439***
	(0.666)	(1.456)	(1.055)	(2.035)	(0.163)	(0.339)	(0.234)	(0.549)
*ChainL*SkillAvg*		-0.033		-0.202		-0.074*		-0.113**
		(0.147)		(0.215)		(0.043)		(0.056)
Importr	-0.249	-0.242	0.439	0.429	-0.379***	-0.381***	0.041	0.029
	(0.438)	(0.433)	(0.547)	(0.548)	(0.134)	(0.134)	(0.153)	(0.153)
*Kx*Ki*			0.301***	0.312***			0.206***	0.210***
			(0.061)	(0.059)			(0.016)	(0.016)
*SKx*Ski*			0.355**	0.331**			0.555***	0.542***
			(0.160)	(0.167)			(0.036)	(0.036)
Distance	-0.458***	-0.458***	-0.435***	-0.435***	-1.596***	-1.596***	-1.591***	-1.591***
	(0.034)	(0.034)	(0.033)	(0.033)	(0.038)	(0.038)	(0.039)	(0.039)
Contiguous	0.454***	0.454***	0.436***	0.436***	0.146	0.146	0.101	0.101
	(0.065)	(0.065)	(0.065)	(0.065)	(0.140)	(0.140)	(0.141)	(0.141)
Legalsystem	0.083*	0.083*	0.111**	0.111**	0.495***	0.495***	0.505***	0.505***
	(0.045)	(0.045)	(0.043)	(0.043)	(0.040)	(0.040)	(0.040)	(0.040)
Colonial	0.415***	0.415***	0.342***	0.342***	1.378***	1.378***	1.369***	1.369***
	(0.106)	(0.106)	(0.107)	(0.107)	(0.129)	(0.129)	(0.131)	(0.131)
Language	0.061	0.061	0.070	0.070	0.267***	0.267***	0.283***	0.283***
	(0.075)	(0.075)	(0.075)	(0.075)	(0.071)	(0.071)	(0.072)	(0.072)
Religion	0.494***	0.494***	0.606***	0.606***	0.663***	0.664***	0.689***	0.689***
	(0.117)	(0.117)	(0.118)	(0.118)	(0.089)	(0.089)	(0.090)	(0.090)
WTO	0.159	0.159	-0.412	-0.412	1.030*	1.030*	0.969*	0.969*
	(0.546)	(0.546)	(0.519)	(0.519)	(0.554)	(0.554)	(0.552)	(0.552)
RTA	0.982***	0.982***	1.029***	1.029***	-0.123	-0.122	-0.091	-0.091
	(0.085)	(0.085)	(0.084)	(0.084)	(0.078)	(0.078)	(0.079)	(0.079)
CU	-0.215**	-0.215**	-0.237**	-0.237**	-1.422***	-1.422***	-1.375***	-1.375***
	(0.109)	(0.109)	(0.109)	(0.109)	(0.229)	(0.229)	(0.236)	(0.236)
Islands	1.536***	1.536***	-5.318***	-5.318***	-4.854***	-4.854***	-4.873***	-4.873***
	(0.153)	(0.153)	(0.723)	(0.723)	(0.360)	(0.360)	(0.360)	(0.360)
Landlocked	-2.646***	-2.646***	-3.172***	-3.172***	-4.196***	-4.196***	-4.261***	-4.261***
	(0.425)	(0.425)	(0.416)	(0.416)	(0.625)	(0.625)	(0.622)	(0.622)
SkillDisp=SkillDispGini								
ChainL	1.587***	0.691	1.348**	-4.657**	2.631***	2.073***	2.536***	2.034***
	(0.521)	(1.587)	(0.638)	(2.055)	(0.144)	(0.424)	(0.168)	(0.469)
*ChainL*SkillDisp*	-5.036***	-4.018*	-2.547	4.626*	-4.233***	-3.642***	-2.465***	-1.922***
	(1.381)	(2.378)	(1.877)	(2.588)	(0.343)	(0.544)	(0.444)	(0.644)
*ChainL*SkillAvg*		0.069		0.453***		0.044		0.039
		(0.115)		(0.161)		(0.032)		(0.034)

SkillDisp=SkillDispNonMid								
ChainL	2.943***	1.061	2.716***	0.494	2.434***	0.281	1.387***	-1.237***
	(0.704)	(0.853)	(0.589)	(1.031)	(0.157)	(0.303)	(0.170)	(0.345)
*ChainL*SkillDisp*	-8.709***	-7.270***	-6.657***	-5.615***	-3.679***	-1.566***	1.193**	3.480***
	(1.874)	(1.747)	(1.563)	(1.352)	(0.414)	(0.474)	(0.463)	(0.526)
*ChainL*SkillAvg*		0.169**		0.214**		0.176***		0.219***
		(0.071)		(0.105)		(0.024)		(0.027)
Observations	167058	167058	148496	148496	165782	165782	147742	147742

Notes: FE1 (exporter FE, importer FE, and industry FE) are also controlled for. Although not reported, in the regressions using *SkillDispGini* and *SkillDispNonMid*, the same control variables as those when using *SkillDispCV* are included. The dependent variable is exports (million USD) from exporter x to importer m in industry i for the PPML regression. As for the EK Tobit estimations, see footnote 19 Robust standard errors, clustered by exporter-importer pair are reported in parentheses. ***p<0.01, **p<0.05, *p<0.1.

Table 6.2 Endogeneity Corrected Estimates

	(1)	(2)	(3)	(4)
	2SLS		EK Tobit with CF	
	All	Manu.	All	Manu.
SkillDisp = SkillDispCV				
*ChainLUS*SkillDisp*	-5.937***	-4.621***	-8.745***	-6.801***
	(0.804)	(0.935)	(0.899)	(0.968)
*ChainLUS*SkillAvg*	-0.377***	-0.327***	-0.556***	-0.487***
	(0.087)	(0.099)	(0.097)	(0.103)
Importr	-1.708***	-1.778***	-2.097***	-2.215***
	(0.096)	(0.100)	(0.088)	(0.092)
*Kx*Ki*		0.111***		0.189***
		(0.016)		(0.015)
*SKx*Ski*		0.337***		0.518***
		(0.038)		(0.041)
Distance	-0.963***	-1.001***	-1.615***	-1.612***
	(0.023)	(0.024)	(0.013)	(0.014)
Contiguous	0.466***	0.426***	0.061	0.013
	(0.075)	(0.078)	(0.041)	(0.046)
Legalsystem	0.247***	0.262***	0.484***	0.493***
	(0.026)	(0.027)	(0.018)	(0.016)
Colonial	0.674***	0.684***	1.335***	1.322***
	(0.075)	(0.079)	(0.047)	(0.045)
Language	0.215***	0.249***	0.327***	0.346***
	(0.042)	(0.043)	(0.029)	(0.029)
Religion	0.398***	0.443***	0.611***	0.632***
	(0.057)	(0.059)	(0.042)	(0.043)
WTO	0.438	0.390	1.054***	0.993***
	(0.394)	(0.437)	(0.194)	(0.191)
RTA	0.228***	0.239***	-0.105***	-0.077***
	(0.047)	(0.049)	(0.033)	(0.030)
CU	0.228**	0.221**	-1.038***	-0.977***
	(0.103)	(0.110)	(0.072)	(0.078)
Islands	-2.780***	-2.906***	-4.685***	-4.706***
	(0.322)	(0.347)	(0.121)	(0.129)
Landlocked	-2.449***	-2.665***	-4.066***	-4.136***
	(0.459)	(0.496)	(0.231)	(0.223)
v1			0.820***	0.683***
			(0.136)	(0.131)
v2			9.974***	7.947***
			(1.153)	(1.128)
Endogeneity test stat.	73.592	21.365	97.920	47.880
p-value	0.000	0.000	0.000	0.000
Weak IV test stat.	527.303	536.048	2038.081	2140.641

SkillDisp = SkillDispGini				
*ChainLUS*SkillDisp*	-9.764***	-5.141***	-12.788***	-9.643***
	(1.096)	(1.107)	(1.063)	(1.184)
*ChainLUS*SkillAvg*	-0.225***	-0.099*	-0.306***	-0.280***
	(0.057)	(0.057)	(0.057)	(0.065)
Endogeneity test stat.	112.867	16.084	163.459	49.378
p-value	0.000	0.000	0.000	0.000
Weak IV test stat.	564.433	634.376	1471.482	1546.562
SkillDisp = SkillDispNonMid				
*ChainLUS*SkillDisp*	-23.341***	-4.465	-33.238***	-13.024***
	(3.823)	(3.708)	(2.924)	(3.264)
*ChainLUS*SkillAvg*	-0.595***	-0.025	-1.014***	-0.337**
	(0.140)	(0.136)	(0.121)	(0.134)
Endogeneity test stat.	56.751	6.069	123.833	23.730
p-value	0.000	0.048	0.000	0.000
Weak IV test stat.	50.006	45.573	149.089	147.001
Observations	64806	59011	162924	145195

Notes: FE1 (exporter FE, importer FE, and industry FE) are also controlled for. Although not reported, in the regressions using *SkillDispGini* and *SkillDispNonMid*, the same control variables as those when using *SkillDispCV* are included. "EK Tobit with CF" stands for EK Tobit with control function method. *v1* and *v2* are the residuals from the two first-stage regressions (see Section 6.2). For more details on the estimation method, see Section 6.2. *SkillDisp*, and *SkillAvg* are instrumented respectively with the corresponding *SkillDisp* measure in 1950, and the *SkillAvg* in 1950. The dependent variable is the logarithm of exports from exporter x to importer m in industry i for the two-stage least squares (2SLS) regression. As for the EK Tobit estimations, see footnote 19. Robust standard errors, clustered by exporter-importer pair in case of 2SLS regression, and bootstrap standard errors based on 300 replications in case of EK Tobit regression are reported in parentheses. The endogeneity (chi-squared) test statistic in the 2SLS regression stands for the difference of the two Sargan-Hansen statistics: one for the equation where the potential endogenous regressors are treated as endogenous, and one for the equation (null hypothesis) where they are treated as exogenous. The endogeneity test in the EK Tobit regression is the Wald test on the null hypothesis that the coefficients of all the three residuals from the first-stage regressions are jointly zero in the second-stage regression. The test statistic for weak instruments is the Kleibergen-Paap Wald rk F Statistic (see Baum *et al*. 2007). The weak instrument test for EK Tobit is based on the 2SLS regression of the second dependent variable. ***p<0.01, **p<0.05, *p<0.1.

Table 6.3 Effect of One Standard Deviation Increase in *SkillDisp* on the *ChainL*'s Impact on Export (In Terms of Percent Change)

	All		Manufacturing	
Table 5.1 OLS estimates excluding zero export				
	col. (3)	col. (4)	col. (3)	col. (4)
SkillDispCV	-18.4	-17.2	-15.2	-18.0
SkillDispGini	-18.1	-17.2	-8.5	-8.8
SkillDispNonMid	-18.4	-17.4	19.8	22.3
Table 6.1 Selection-corrected estimates				
	PPML	EK Tobit	PPML	EK Tobit
	col. (2)	col. (6)	col. (4)	col. (8)
SkillDispCV	-46.2	-44.5	-65.4	-40.4
SkillDispGini	-38.3	-35.5	74.5	-20.6
SkillDispNonMid	-53.5	-15.2	-44.6	44.3
Table 6.2 Endogeneity-corrected estimates				
	2SLS	EK Tobit/CF	2SLS	EK Tobit/CF
	col. (1)	col. (3)	col. (2)	col. (4)
SkillDispCV	-71.6	-84.3	-62.4	-76.3
SkillDispGini	-69.1	-78.5	-46.1	-68.7
SkillDispNonMid	-91.4	-97.0	-37.5	-74.6

Notes: The above statistic is computed as [exp($\hat{\beta}_2$ **SD*)-1]*100(%), where *SD* is one standard deviation of *SkillDisp* (0.212 for *SkillDispCV*, 0.120 for *SkillDispGini*, and 0.105 for *SkillDispNonMid*). $\hat{\beta}_2$, which is the estimated coefficient of *ChainL*SkillDisp* (or *ChinLUS*SkillDisp*), is taken from the corresponding tables. In all specifications, *ChianL*SkillAvg* (or *ChainLUS*SkillAvg*) is controlled for. Highlighted cells are those using insignificant estimates.

Table 6.4 Endogeneity Corrected Estimates with Alternative Factors

(a) 2SLS estimates

	(1)	(2)	(3)	(4)	(5)	(6)
		All			Manufacturing	
AltFactor=	*Eleloss*	*Roadp*	*RoLaw*	*Eleloss*	*Roadp*	*RoLaw*
SkillDisp = SkillDispCV						
*ChainLUS*SkillDisp*	-5.833***	-2.703***	-5.617***	-5.480***	-3.421***	-4.723***
	(0.816)	(0.787)	(0.774)	(0.965)	(0.930)	(0.933)
*ChainLUS*SkillAvg*	-0.391***	-0.164*	-0.538***	-0.343***	-0.202**	-0.298***
	(0.086)	(0.091)	(0.088)	(0.098)	(0.103)	(0.101)
*ChainLUS*AltFactor*	-1.284	1.268***	0.566***	8.015***	-1.323***	-0.122**
	(1.041)	(0.133)	(0.053)	(1.213)	(0.144)	(0.058)
Endogeneity test stat.	62.264	7.181	39.471	23.605	14.915	28.873
p-value	0.000	0.028	0.000	0.000	0.001	0.000
Weak IV test stat.	565.676	530.795	652.287	585.684	522.390	646.744
SkillDisp = SkillDispGini						
*ChainLUS*SkillDisp*	-9.162***	-5.210***	-8.029***	-5.015***	-4.019***	-5.351***
	(1.086)	(1.030)	(0.990)	(1.113)	(1.103)	(1.078)
*ChainLUS*SkillAvg*	-0.235***	-0.164***	-0.369***	-0.038	-0.061	-0.084
	(0.056)	(0.063)	(0.058)	(0.057)	(0.066)	(0.060)
*ChainLUS*AltFactor*	-2.034*	1.501***	0.653***	6.934***	-1.153***	-0.075
	(1.042)	(0.141)	(0.056)	(1.178)	(0.150)	(0.058)
Endogeneity test stat.	97.911	27.985	81.071	15.793	14.705	20.214
p-value	0.000	0.000	0.000	0.000	0.001	0.000
Weak IV test stat.	545.348	1174.502	802.174	595.865	1218.628	834.991
SkillDisp = SkillDispNonMid						
*ChainLUS*SkillDisp*	-20.401***	-30.005***	-17.020***	-14.280***	-25.157*	-8.586*
	(3.189)	(9.320)	(4.073)	(3.547)	(13.100)	(4.501)
*ChainLUS*SkillAvg*	-0.544***	-0.997***	-0.497***	-0.320**	-0.788	-0.100
	(0.118)	(0.357)	(0.133)	(0.131)	(0.500)	(0.151)
*ChainLUS*AltFactor*	-3.681***	-0.352	0.359***	5.913***	-2.731***	-0.215***
	(1.117)	(0.494)	(0.078)	(1.218)	(0.715)	(0.078)
Endogeneity test stat.	50.693	19.070	28.985	36.059	6.814	7.015
p-value	0.000	0.000	0.000	0.000	0.033	0.030
Weak IV test stat.	71.830	11.699	29.170	63.467	6.205	25.989
Observations	63911	58370	64806	58169	53063	59011

(b) EK Tobit estimates with control function approach

	(1)	(2)	(3)	(4)	(5)	(6)
		All			Manufacturing	
AltFactor=	*Eleloss*	*Roadp*	*RoLaw*	*Eleloss*	*Roadp*	*RoLaw*
SkillDisp = SkillDispCV						
*ChainLUS*SkillDisp*	-8.379***	-6.831***	-8.506***	-7.084***	-5.880***	-6.950***
	(0.912)	(1.031)	(0.915)	(0.970)	(1.035)	(0.980)
*ChainLUS*SkillAvg*	-0.634***	-0.529***	-0.824***	-0.501***	-0.403***	-0.531***
	(0.098)	(0.113)	(0.101)	(0.104)	(0.112)	(0.111)
*ChainLUS*AltFactor*	-8.543***	1.938***	0.890***	3.475***	-1.118***	0.102
	(1.046)	(0.163)	(0.066)	(1.116)	(0.169)	(0.073)
Endogeneity test stat.	89.182	23.814	67.549	43.064	34.350	49.285
p-value	0.000	0.000	0.000	0.000	0.000	0.000
Weak IV test stat.	1905.850	1879.568	2224.293	1974.364	1931.005	2315.398
SkillDisp = SkillDispGini						
*ChainLUS*SkillDisp*	-12.311***	-11.035***	-10.490***	-8.972***	-9.068***	-9.322***
	(1.058)	(1.185)	(1.049)	(1.180)	(1.326)	(1.163)
*ChainLUS*SkillAvg*	-0.434***	-0.433***	-0.499***	-0.250***	-0.308***	-0.314***
	(0.060)	(0.071)	(0.060)	(0.067)	(0.079)	(0.069)
*ChainLUS*AltFactor*	-11.278***	2.464***	0.970***	1.926*	-0.686***	0.163**
	(1.022)	(0.170)	(0.065)	(1.095)	(0.175)	(0.072)
Endogeneity test stat.	144.923	54.641	122.730	39.513	41.130	48.794
p-value	0.000	0.000	0.000	0.000	0.000	0.000
Weak IV test stat.	1861.351	1954.377	2123.706	1942.523	1990.076	2154.154
SkillDisp = SkillDispNonMid						
*ChainLUS*SkillDisp*	-23.304***	-35.820***	-23.442***	-16.063***	-23.940***	-14.246***
	(2.472)	(4.413)	(3.431)	(2.700)	(5.025)	(3.609)
*ChainLUS*SkillAvg*	-0.785***	-1.258***	-0.809***	-0.484***	-0.761***	-0.362***
	(0.106)	(0.175)	(0.129)	(0.118)	(0.203)	(0.139)
*ChainLUS*AltFactor*	-13.582***	-0.182	0.591***	-0.906	-2.572***	-0.075
	(1.118)	(0.307)	(0.076)	(1.270)	(0.330)	(0.078)
Endogeneity test stat.	85.206	69.626	72.095	61.123	31.235	23.975
p-value	0.000	0.000	0.000	0.000	0.000	0.000
Weak IV test stat.	268.751	79.770	104.255	264.588	77.452	102.560
Observations	160066	148634	162924	142648	132460	145195

Notes: The only difference between this table and Table 6.2 is that this table additionally controls for *ChainLUS*AltFactor*, where *AltFactor* is either *Eleloss*, *Roadp*, or *RoLaw*. For more detail, see Table 6.2 and Section 6.3. ****p*<0.01, ***p*<0.05, **p*<0.1.

Appendix Table 1. *SkillDisp* Indices and *SkillAvg* by 58 Exporters

Rank	*SkillDispCV*		*SkillDispGini*		*SkillDispNonMid*		*SkillAvg*	
1	US	0.207	US	0.103	Slovenia	0.133	US	12.59
2	Slovenia	0.236	Russia	0.126	Poland	0.138	Czech Rep.	11.99
3	Czech Rep.	0.241	Canada	0.135	Czech Rep.	0.154	New Zealand	11.68
4	Canada	0.284	Latvia	0.143	Hungary	0.162	Slovakia	11.24
5	Australia	0.285	Australia	0.153	Italy	0.173	Australia	11.20
6	Sweden	0.301	Denmark	0.175	Austria	0.178	Sweden	11.18
7	Slovakia	0.311	Lithuania	0.181	Slovakia	0.183	Israel	11.07
8	Netherlands	0.314	Netherlands	0.183	Croatia	0.185	Slovenia	11.05
9	Hungary	0.319	Slovenia	0.192	Latvia	0.199	Norway	10.72
10	Denmark	0.324	Sweden	0.194	Norway	0.209	Canada	10.71
11	Norway	0.328	Japan	0.203	Estonia	0.210	Netherlands	10.56
12	Japan	0.334	Estonia	0.219	Netherlands	0.228	Japan	10.51
13	Estonia	0.338	Romania	0.224	Denmark	0.229	South Korea	10.49
14	Romania	0.346	Norway	0.226	Lithuania	0.233	Hungary	10.41
15	Latvia	0.355	Hungary	0.239	Sweden	0.251	Ireland	10.39
16	New Zealand	0.357	Austria	0.242	Malta	0.252	Estonia	10.38
17	Poland	0.370	South Korea	0.247	Romania	0.259	Denmark	10.04
18	Austria	0.396	Philippines	0.249	Iceland	0.262	Belgium-Lux	9.97
19	Israel	0.397	Poland	0.250	Bulgaria	0.263	Russia	9.94
20	Russia	0.397	Belgium-Lux	0.258	Greece	0.281	Poland	9.86
21	South Korea	0.400	Taiwan	0.263	Germany	0.288	Switzerland	9.83
22	Ireland	0.403	New Zealand	0.264	Argentina	0.289	Romania	9.70
23	Belgium-Lux	0.411	Slovakia	0.270	France	0.293	Cyprus	9.60
24	Lithuania	0.416	Czech Rep.	0.270	South Korea	0.296	Germany	9.44
25	Iceland	0.425	South Africa	0.273	UK	0.299	UK	9.44
26	UK	0.439	Finland	0.275	Japan	0.299	Malta	9.36
27	Croatia	0.448	UK	0.280	Portugal	0.305	Iceland	9.23
28	Germany	0.453	Argentina	0.280	Finland	0.306	Lithuania	9.17
29	Bulgaria	0.461	France	0.283	Ireland	0.307	Croatia	9.08
30	Finland	0.466	Ireland	0.287	South Africa	0.314	Latvia	8.99
31	Argentina	0.466	Italy	0.288	Canada	0.314	Bulgaria	8.86
32	France	0.471	Iceland	0.288	Belgium-Lux	0.319	France	8.82
33	Cyprus	0.473	Chile	0.291	Malaysia	0.323	Austria	8.81
34	South Africa	0.474	Bulgaria	0.299	China	0.328	Chile	8.78
35	Chile	0.482	Israel	0.300	Australia	0.333	Taiwan	8.77
36	Italy	0.501	Croatia	0.302	Cyprus	0.338	Greece	8.73
37	Taiwan	0.503	Cyprus	0.317	Taiwan	0.349	Argentina	8.64
38	Malta	0.504	Costa Rica	0.319	Switzerland	0.356	Finland	8.59
39	Switzerland	0.509	Germany	0.322	Costa Rica	0.364	Malaysia	8.39
40	Greece	0.533	Viet Nam	0.356	Israel	0.366	South Africa	8.29
41	Malaysia	0.561	China	0.361	Chile	0.385	Italy	8.27
42	Costa Rica	0.586	Thailand	0.371	Singapore	0.389	Spain	8.11
43	Philippines	0.602	Malaysia	0.373	Turkey	0.392	Singapore	7.98

44	Singapore	0.608	Greece	0.379	Spain	0.410	Philippines	7.56
45	Mexico	0.647	Colombia	0.390	Mexico	0.414	Costa Rica	7.42
46	Spain	0.660	Mexico	0.398	Viet Nam	0.422	Mexico	7.20
47	Portugal	0.672	Malta	0.399	New Zealand	0.442	Portugal	6.69
48	Colombia	0.706	Singapore	0.399	Philippines	0.444	Colombia	6.47
49	Thailand	0.724	Switzerland	0.405	Colombia	0.454	Saudi Arabia	6.34
50	China	0.731	Portugal	0.416	US	0.457	China	6.32
51	Viet Nam	0.765	Brazil	0.464	Russia	0.460	Brazil	5.58
52	Brazil	0.804	Spain	0.466	Brazil	0.462	Thailand	5.50
53	Saudi Arabia	0.861	Turkey	0.542	Saudi Arabia	0.522	Turkey	5.44
54	Turkey	0.916	Saudi Arabia	0.543	Indonesia	0.527	Tunisia	5.07
55	Indonesia	0.997	Indonesia	0.611	Thailand	0.552	Viet Nam	4.85
56	Tunisia	1.038	Cambodia	0.655	Tunisia	0.558	Indonesia	4.62
57	Cambodia	1.187	Tunisia	0.662	India	0.576	India	4.12
58	India	1.392	India	0.692	Cambodia	0.723	Cambodia	3.27

Notes: Lower rank corresponds to lower skill dispersion or higher average skill of the economy. For the construction of indices, see Section 4.2.
Source: Constructed from Barro and Lee (2013).

Appendix Table 2. Variable Description and Summary Statistics

Variable	Mean	Std. Dev.	Description
ln(Export)	8.241	2.342	Logarithm of export value (1000 nominal USD) in exporter-importer-industry cell in 2000 (See Section 4.1).
ChainL	1.818	0.283	Industry-exporter-specific length of domestic production chains (See Section 4.3).
SkillDisp -CV	0.502	0.212	Skill dispersion measure (CV) of exporting country (hereafter, exporter) in 1995 (See Section 4.2).
SkillDisp -Gini	0.306	0.120	Skill dispersion measure (Gini coefficient) of exporter in 1995 (See Section 4.2).
SKillDisp -NonMid	0.321	0.105	Skill dispersion measure (one minus the ratio of semi-skilled populations over age 15, who have completed at least primary education but have not received any tertiary education) of exporter in 1995 (See Section 4.2).
SkillAvg	8.897	1.948	Average skill (average years of schooling of the populations over age 15) of exporter (See Section 4.2).
Importr	0.288	0.153	Ratio of the imported inputs to total input value in each industry. *Source*: OECD (2015)
Kx	4.657	0.806	Capital intensity of exporter = logarithm of [ratio of capital stock at current PPPs (in thousand 2005USD) to total employment]. *Source*: Penn World Table, version 8.1 (Feenstra *et al*. 2015)
Ki	4.362	0.715	Capital intensity of industry (only available for manufacturing) = logarithm of [1995-1999 average ratio of real capital stock (in thousand USD) to total employment of US industries]. Matching of original US SIC 1987 industry code with the 18 industries based on ISIC Rev.3 is based on the concordance table provided by the United Nations (http://unstats.un.org/unsd/cr/registry/regdnld.asp?Lg=1). *Source*: Becker *et al*. (2014).
SKx	8.913	1.941	Skill intensity of exporter = *SkillAvg*.
SKi	0.285	0.096	Skill intensity of industry (only available for manufacturing) = 1995-1999 average ratio of non-production workers to total employment of US industries. *Source*: Becker *et al*. (2014).
Distance	8.383	0.981	Logarithm of the distance between the capital cities of exporter and importing country (hereafter, importer). *Source*: CEPII's distance data (see Mayer and Zignago, 2006).
Contiguous	0.046	0.209	Dummy = 1 if exporter and importer are contiguous. *Source*: CEPII's distance data (see Mayer and Zignago, 2006).
Legalsystem	0.289	0.453	Dummy = 1 if exporter and importer share the same legal origin. *Source*: Helpman *et al*. (2008).
Colonial	0.036	0.187	Dummy = 1 if importer ever colonized exporter or vice versa.

			Source: Helpman *et al*. (2008).
Language	0.124	0.330	Dummy = 1 if a language is spoken by at least 9% of the population in both exporter and importer. *Source*: CEPII's distance data (see Mayer and Zignago, 2006).
Religion	0.170	0.222	The degree of shared religion between exporter and importer. It is constructed as follows by applying the method of Helpman *et al*. (2008). Religion_*xm* = $\sum_k$ (% religion_*k* in exporter * % religion_*k* in importer), where % religion_*k* indicates percentage of population who are adherent to religion *k*. There are nine religions (Catholic, Protestant, other Christian, Orthodox, Muslim, Hindu, Buddhist, Other Eastern religions, and Jewish). *Source*: Barro and McCleary (2005).
WTO	1.689	0.490	Number of WTO members in exporter-importer pair *Source*: WTO website.
RTA	0.194	0.387	Dummy = 1 if exporter and importer belong to a common regional trade arrangement. *Source*: Head, Mayer, and Ries (2010).
CU	0.017	0.106	Dummy = 1 if exporter and importer share a currency. *Source*: Head, Mayer, and Ries (2010).
Islands	0.270	0.486	Number of islands in in exporter-importer pair. *Source*: Helpman et al. (2008)
Landlocked	0.218	0.440	Number of landlocked countries in exporter-importer pair. *Source*: Helpman et al. (2008)
Eleloss	0.088	0.046	Exporter-specific electric power transmission and distribution losses (ratio to output), averaged over non-missing years 1995-1999. *Source*: World Bank (2014).
Roadp	0.712	0.293	Exporter-specific ratio of paved roads to total roads, averaged over non-missing years 1995-1999. Only for Lithuania and US, year 2000 data are used, because 1995-1999 data are not available. *Source*: World Bank (2014).
RoLaw	0.849	0.848	Exporter-specific rule of law index (ranging from -2.5 [weak] to 2.5 [strong]) for the year 2000. *Source*: The Worldwide Governance Indicators, 2016 Update (www.govindicators.org) (see Kaufmann et al. [2010] for detail).

Notes: Summary statistics are based on column (3) of Table 5.1 (number of observation is 65858) except for *Kx* (number of observations is 59946), *Ki* (59946), *SKx* (59946), *SKi* (59946), *Eleloss* (63911), *Roadp* (58370), and *RoLaw* (64806). Note that the statistics on *Export* are based on non-zero export.

Appendix Table 3. Length of Production Chains (*ChainL*) for All Industries

Industry	*ChainL* Mean	*ChainL* Std. Dev.
Private households with employed persons	1.000	0.000
Education	1.303	0.176
Real estate activities	1.332	0.170
Public administration and defense; compulsory social security	1.457	0.180
Health and social work	1.475	0.221
Mining and quarrying	*1.490*	*0.214*
Financial intermediation	1.518	0.160
Wholesale and retail trade; repairs	1.545	0.179
Renting of machinery and equipment	1.561	0.241
R&D and other business activities	1.585	0.206
Coke, refined petroleum products and nuclear fuel	1.591	0.310
Post and telecommunications	1.591	0.214
Computer and related activities	1.613	0.260
Other community, social and personal services	1.614	0.188
Electricity, gas and water supply	1.624	0.298
Computer, Electronic and optical equipment	1.635	0.226
Agriculture, hunting, forestry and fishing	*1.642*	*0.192*
Transport and storage	1.685	0.181
Other non-metallic mineral products	1.694	0.221
Machinery and equipment, nec	1.715	0.270
Other transport equipment	1.722	0.277
Electrical machinery and apparatus, nec	1.736	0.275
Fabricated metal products	1.737	0.244
Hotels and restaurants	1.763	0.201
Textiles, textile products, leather and footwear	1.770	0.258
Chemicals and chemical products	1.778	0.267
Rubber and plastics products	1.781	0.255
Construction	1.784	0.242
Basic metals	1.788	0.310
Motor vehicles, trailers and semi-trailers	1.792	0.345
Manufacturing nec; recycling	1.792	0.219
Pulp, paper, paper products, printing and publishing	1.822	0.220
Wood and products of wood and cork	1.900	0.213
Food products, beverages and tobacco	2.061	0.199

Notes: Highlighted industries in *Italic* are primary industries. The remaining highlighted industries are manufacturing industries. The statistics, which are arranged in ascending order of *ChainL*, are unweighted average for 58 exporters. They are different from those in Table 4.3 due to unweighting.
Source: OECD (2015).

(This page is intentionally left blank)

Chapter 4

Delegation to Workers across Countries and Industries: Social Capital and Coordination Needs Matter*†

Abstract

The degree of delegating authority to non-managerial and non-supervisory workers substantially varies across countries and industries. By examining worker-level data from 14 countries, I empirically explain this variation by region-specific social capital that proxies workers' degree of self-centeredness and the industry-specific need for coordination. The empirical results of this study confirm the theoretical predictions by Alonso *et al*. (2008) for the first time: the negative association between coordination needs and decentralization is mitigated in regions with lower self-centeredness of workers. In particular, when self-centeredness of workers (respectively, need for coordination) is very low, the degree of delegation is always high regardless of the level of the need for coordination (self-centeredness of workers). Positive associations between delegation and its benefits, including job satisfaction, wages (proxy for higher productivity), and skill upgrading of workers, are also found. These results imply that people's degree of self-centeredness affects a country's economic development patterns by changing the degree of decentralization and its benefits.

JEL Classification: L22, L23, Z13

Keywords: Coordination, Decentralization, Delegation, PIAAC, Social Capital, Trust

---

* Previous version of this chapter was released as the IDE discussion paper No. 620 (http://www.ide.go.jp/English/Publish/Download/Dp/620.html).

† I thank Yuji Genda, Hideaki Goto, Daiji Kawaguchi, Ayako Kondo, Tomohiro Machikita, Hideo Owan, and all the internal seminar participants at the Institute of Developing Economies, Japan External Trade Organization (IDE-JETRO) and the University of Tokyo. I am grateful for the funding support from the IDE-JETRO.

The main text of Chapter 4 will be disclosed within five years.

For the earlier version of Chapter 4, see the IDE discussion paper No. 620 (http://www.ide.go.jp/English/Publish/Download/Dp/620.html).

Chapter 5

Concluding Remarks

The previous three chapters have empirically shown that the impact of human/social capitals on a certain outcome (wages, export, and delegation to workers) depends on the technological characteristics of a given industry, which is primarily characterized by the scope of production linkages. Such different impacts across industries, in turn, shape the patterns of economic development, such as skill allocation across sectors and a country's comparative advantage. These empirical results contribute to the literature by shedding light on new nexuses among human/social capitals, industry technology, and economic development.

In this chapter, I conclude the dissertation by highlighting several areas for future research. First, the empirical analyses in the previous three chapters are more or less reduced form. More micro-evidence that directly tests the hypothesized mechanisms is required for the future. For example, more direct evidence on (i) whether quality deteriorates along the production chains (Chapter 2); (ii) whether comparative advantage is affected by the country's skill-sorting pattern (Chapter 3); and (iii) whether horizontal communication and coordination among workers increase when both social capital and coordination needs are high (Chapter 4) would further strengthen the arguments of this dissertation.

Second, as mentioned in Chapter 1, all three studies share the same empirical strategy and focus on the coefficient of an interaction term, which is the term $\beta_3 Z_{ij} * X_i$ in equation (2.1) in Chapter 1. Although this strategy is frequently used in macroeconomics and international trade literature (Ciccone and Papaioannou 2016), it has the following limitations: (i) It is difficult to control for various factors other than

Z_{ij} and X_i, which may change the coefficient β_3. This is because a multicollinearity problem arises as the number of other factors (*O*), which are also controlled by interaction terms ($O * X_i$ or $O * Z_{ij}$), increases. (ii) When cross-country or cross-region variations are examined, as in Chapters 3 and 4, the estimating equation usually assumes that the effects of various control variables are the same regardless of countries/regions. For example, in Chapter 4, it is implicitly assumed that the effect of education on delegation is the same across all 14 countries, although such a strong assumption may be inevitable without a huge sample size. In sum, alternative estimation strategies that overcome the problems (i) and (ii) are necessary.

Third, as summarized in Chapter 1, I use the column sum of the Leontief inverse coefficient ($Leon_j$), which is computed from the input–output tables, as the primary industry-specific measure for the scope of production linkages. Constructing and examining alternative measures would certainly strengthen the results obtained in this dissertation.

Finally, this dissertation primarily assumes that the technological characteristics of industries are exogenous. The possible endogeneity is minimized by alternatively using the industry characteristics of a benchmark country (i.e., the United States) in robustness checks (Chapters 3 and 4). However, it is highly likely that country characteristics affect a firm's adoption choice of industry technology. For example, firms operating in countries with less-advanced technologies may shorten the length of domestic production chains by importing higher-quality intermediate inputs. They may also mitigate the quality deterioration along the production chains by introducing foreign capital and knowledge and by simplifying the production process through modular systems or 3D printers. Endogenizing the industry technology and examining the two-way interactions between country characteristics and industry technology is an interesting arena for future study.

References

Ciccone, Antonio, and Elias Papaioannou. 2016. “Estimating Cross-Industry Cross-Country Interaction Models Using Benchmark Industry Characteristics.” NBER Working Paper 22368.