RIETI Discussion Paper Series 18-E-083

Skill Transference and International Migration: A theoretical analysis on skilled migration to the Anglosphere

NAKAGAWA Mariko
University of Tokyo

Research Institute of Economy, Trade & Industry, IAA
The Research Institute of Economy, Trade and Industry
https://www.rieti.go.jp/en/

RIETI Discussion Paper Series 18-E-083
December 2018

Skill Transference and International Migration: A theoretical analysis on skilled migration to the Anglosphere[*]

NAKAGAWA Mariko
Center for Spatial Information Science, the University of Tokyo

Abstract

In this paper, we analyze how skill transference from an origin to destination country, captured by lower productivity at the destination caused by differences in language use, affects international migration by skilled workers, using a multi-country NEG model proposed by Gasper et al. (2017). Specifically, our interest is to explain how less frictional countries in terms of linguistic communication such as those in the Anglosphere (English-speaking countries) attract more highly-skilled international migrants. The analysis based on asymmetric skill transference among countries, in which the world is divided into two groups, Anglosphere (English-speaking countries) and non-Anglosphere (non-English-speaking countries), finds that countries in the Anglosphere are more likely to be the industrial core attracting all skilled (and imperfectly mobile) workers than countries in the non-Anglosphere. Also, we find that less frictional migration from the non-Anglosphere to the Anglosphere always accelerates industrial agglomeration in the Anglosphere core country, while both less frictional migration within the Anglosphere and expanding the Anglosphere (an increase in the number of countries constituting the Anglosphere) do not always accelerate industrial agglomeration in the Anglosphere due to the market crowding effect.

Keywords: Skill transference, International migration, Anglosphere, Frictional cost, Language difference
JEL classification: F22, R13, F66

[*]This study is conducted as a part of the Project "Spatial Economic Analysis on Intracity Economic Activities and Interregional Economic Activities" undertaken at the Research Institute of Economy, Trade and Industry (RIETI). The author is grateful for helpful comments and suggestions by Takatoshi Tabuchi, Hikaru Ogawa, Tetsuji Okazaki, Takaaki Takahashi, Dan Sasaki, Yasuhiro Sato, Toshihiro Okubo, Ryosuke Okamoto, Shota Fujishima, Tadashi Morita, Keisuke Kawata, Jun Oshiro, Yukihiro Kidokoro, Yuichiro Yoshida, Hajime Takatsuka, Kohei Nagamachi, and participants at UEA annual meeting in Dusseldorf, Okinawa University seminar, Urban Economics Workshop at the University of Tokyo, Applied Economic Workshop at Keio University, IDEC seminar at Hiroshima University, Regional Science seminar at Kagawa University, and Discussion Paper seminar participants at RIETI. This study is supported by Grant-in-Aid for Research Activity Start-up (Research project number: 16H06703) and Grant-in-Aid for Young Scientists (Research project number: 18K12764) provided by the Japan Society for the Promotion of Science (JSPS).

†Center for Spatial Information Science, University of Tokyo. E-mail: mnakagawa@csis.u-tokyo.ac.jp. Address: 5-1-5Kashiwanoha, Kashiwa-shi, Chiba, 277-8568, Japan

1 Introduction

When facing a problem on whether to migrate to a foreign country or stay in the home country, what factors affect potential migrants' decisions? If having decided to leave their home country, where should they choose to reside? There may be a lot of elements affecting such decisions of migrants. They may prefer destinations that would offer better circumstances with safer, richer, and higher living standards. At the same time, they may attempt to circumvent difficulties and obstacles which can prevent them from a smooth shift from the life in their origin country to a life in the destination by assimilating to a new society.

Because migrants' decision and life in the destination country are subject to various social and economic aspects, it would be difficult to reach the most decisive point to determine potential migrants' decisions. However, based on Roy's (1951) idea that migrants' decision on whether to migrate or stay and where to migrate is related to utility maximization, how smoothly potential migrants can transfer their origin-accumulated skills to the destination, which directly affects their expected earnings there, should be a crucial concern associated with international migration.

A primarily important observation about what the source of friction accompanied by skill transfer in migration is comes from Chiswick and Miller (2014), in which the authors argue that the set of skills useful in origin countries does not necessarily coincide with useful skills in destination countries. Differences in technology, custom, occupational license, and language are some of the examples which induce difficulty in skill transference for migrants. Cardinal importance of international skill transfer and associated costs is also considered in other contexts, such as task trading (Grossman and Rossi-Hansberg, 2008) and offshoring (Ottaviano et al., 2013), both shedding light on the aspects of frictional costs associated with labor supply in that skills or labor endowments cannot be fully utilized due to differences in culture, operation, and languages. Among these sources of difficulties in skill transfer, this paper especially focus on the frictional cost related to language differences which lowers productivity at the destination countries because it is often that a language dominantly used in the origin is different from that in the destination when migration crosses national borders. Migrants suffering from mismatch in language use between the origin and destination confront an impossibility of fully demonstrating their abilities in the environment of languages different from their mother tongue.[1]

[1]Note that this paper does not exclude the investigation of other obstacles which affect migration decision than language barriers. Rather, we can consider any type of frictional costs lowering productivity at the destination country in the form of frictional costs used in Section 2. As an illustrative example, we highlight language barriers associated

Accordingly, how and to what extent language affects skill transference is investigated and certified in the empirical literature. Belot and Hatton (2012) find that if the language barrier that migrants face is less severe, then the transferability of their human capital to a destination is greater. In relation with language proficiency and earnings at destination, Dustmann and Van Soest (2001, 2002) assert that higher proficiency of language in the destination country increases immigrants' earnings. These empirical literatures support an idea that potential migrants should take into account the linguistic aspect which would affect their expected income at the destination countries so that they tend to avoid frictional costs stemming from being less productive at destination due to language barrier.

As stated above, since international migration is frictional in terms of labor productivity, potential migrants would choose destinations of smoother skill transfer. Then, from a viewpoint of language use, what countries would be candidates of destination which will bring about less frictional skill transfer? One prepotent option may be English-speaking countries (called the Anglosphere) due to its large population of speakers as an acquired language and its worldwide perception.[2] Figure 1 in Appendix A exhibit how the Anglosphere attracts international migrants, by showing the long-term trend of immigrant inflows to the OECD destinations, starting from year 1946, ending with year 2011. Blue (red, respectively) lines indicate immigrant inflows (in thousands) to the Anglosphere (non-Anglosphere, respectively) destinations. Comparing the trends of immigrant inflows of the Anglosphere destinations and the non-Anglosphere destinations suggests that among OECD destinations, countries belonging to the Anglosphere attract more international migrants than non-Anglosphere destinations, which may be a reflection of migrants' tendency of choosing less frictional destinations in view of language difference.

Furthermore, Clark et al. (2007) find that source countries in the Anglosphere tend to exhibit a higher migration rate to the United States compared to other countries, which implies that the migration cost to the United States is lower if a dominant language in the origin country is English. Similarly, as discussed in Adsera and Pytlikova (2015) and Grogger and Hanson (2011), countries

with international migration.

[2]There may be other important candidates of less friction-inducing destinations in terms of language difference. For example many post-colonized countries adopt colonizer's language as official languages, which consist of "language world" such as Spanish-speaking world, French-speaking world and so on. As discussed in Pedersen et al. (2008) and Belot and Ederveen (2012), common languages between origin and destination countries are known to play important roles for immigrants in OECD countries. In a model presented in Section 2, however, this aspect is not dealt with due to the dichotomous classification of the world to the Anglosphere and the non-Anglosphere. In this sense, our model may target only on the international migration between English-speaking countries and other countries whose official language is not world-widely used. Nontheless, we conducted an analysis to take into account the language world and our results indicate that there is some possibility of agglomeration within each language world, although the result is not reported in this paper. Results are available upon request.

in the Anglosphere are inclined to attract more migrants than destinations in the non-Anglosphere, other things being equal. Particularly, Adsera and Pytlikova's (2015) reasoning about the attraction of immigrants to the Anglosphere destinations supports our motivation: migrating to English-speaking destinations may be less friction-inducing in skill transferring even for immigrants whose mother tongue is not English because a large population have learned English as a second language in the elementary or secondary levels of education in their home country.

This tendency of choosing Anglosphere destinations rather than non-Anglosphere destinations may be more apparent in skilled migration. Table 1 in Appendix B reports the share of skilled immigrants (immigrants with a college degree) to the total immigrants to OECD destinations, in which migration patterns are decomposed into four categories, depending on whether origin and destination countries belong to the Anglosphere (countries whose official language is English) or non-Anglosphere (countries whose official language is not English). Figures in Table 1 are shares of "the stock of immigrants with a college degree whose age of entry is 22 and over to OECD destinations" to "the stock of immigrants whose age of entry is 22 and over to OECD destinations."[3] By comparing four categories of the origin-destination pairs, high-skilled (with a college degree) immigrant stock share to the total immigrant stock is highest for the pair of an Anglosphere origin and a non-Anglosphere destination both for year 2000 and 1990. The second highest is the pair of a non-Anglosphere origin and an Anglosphere destination, and the lowest is that of a non-Anglosphere origin and a non-Anglosphere destination. These may indicate that high-skilled migrants tend to migrate within the Anglosphere because this type of migration may induce smaller obstacles against skill transfer due to the same language in the origin and destination countries. In addition, Anglosphere destinations are more likely to be chosen by high-skilled migrants than non-Anglosphere destinations, which may reflect smoother skill transference in the case of migration to Anglosphere destinations than non-Anglosphere destinations.

Following these observations, in this paper, we model the tendency of potential skilled migrants to prefer destination countries with less friction in order to circumvent lower productivity or earnings at destination due to mismatch in language use. By adopting an n-country footloose entrepreneur model with a quasi-linear upper tier utility function proposed in Gaspar et al. (2017), and introducing the friction against skill transference associated with international migration to the model, we investigate

[3]It is fair to limit the sample whose age of entry is 22 and over because in many cases, immigrants with college degree tend to be older than 22 years old. In addition, since we focus on the friction of skill transfer caused by language barriers, it is reasonable to restrict the sample to immigrants whose age of entry is 22 and over. Because adapting to environments with different languages from one's mother tongue may be harder for adults than younger immigrants, it is adequate to limit samples to those with the age of entry over 22 when the purpose of showing Table 1 is to illustrate the skilled migration tendency to choose destinations with less friction of skill transfer.

how skilled (and imperfectly mobile) worker distribution in equilibrium is affected by the frictional term expressing less productivity at destination. The analysis finds that countries in the Anglosphere are more likely to be the industrial core attracting all skilled (and imperfectly mobile) workers than countries in the non-Anglosphere in stable equilibrium when the extent of friction associated with migration to the non-Anglosphere is sufficiently high.

In addition, we investigate the impact of a "spread of English communication" on the stable equilibrium characterized by the fully agglomerated configuration in the Anglosphere. With the normal interpretation, the Anglosphere is considered as countries adopting English as an official language. With an extended interpretation that we use in the analysis, by contrast, the Anglosphere consists of countries where English communication is relatively smooth, such as European countries, and accordingly, the non-Anglosphere consists of countries where English communication tends to be difficult, such as some Asian countries. Then, the situation in which English is becoming more and more common corresponds to (i) an expansion of the Anglosphere (a number of countries within the Anglosphere increases, or put differently, smooth English communication is undertaken in more countries), (ii) less frictional migration to the Anglosphere, and (iii) less frictional migration within the Anglosphere. By inspecting the impacts of the spread of English on the sustain condition (stability condition for an equilibrium of industrially agglomeration in the Anglosphere), we find that less frictional migration from the non-Anglosphere to the Anglosphere always accelerates industrial agglomeration in the Anglosphere core country through a channel of less severe friction in terms of skill transfer, while less frictional migration within the Anglosphere and an expansion of the Anglosphere (an increase in the number of countries constituting the Anglosphere) do not necessarily accelerate industrial agglomeration in the Anglosphere due to the market crowding effect.

Finally in this section, we clarify where this paper is positioned in the spatial economics. Merging migration costs and spatial economics as is conducted in this paper itself is not new. Ludema and Wooton (1999) model imperfect mobility of manufacturing workers by adding a variable which expresses living and working preferences stemming from non-pecuniary elements and affects a relative real wage difference. Also, Tabuchi et al. (2017) introduce to a standard model of the new economic geography the migration cost from one region to the other as an additive term. What we do in this paper is to explicitly and directly introduce the migration costs which negatively affect productivity at destination, by focusing on the aspects of obstacles against smooth skill transfer under international migration. Because in international migration skill transference may be far less smooth than regional migration, introducing frictional skill transfer in the model is necessary. However, the primal object of this paper is not the point related to adding the migration cost expressing lower productivity at destination but analyzing a world which is divided into two heterogeneous groups in terms of freeness of migration, the Anglosphere and the non-Anglosphere. The asymmetry in migration costs is expressed by the less frictional migration to the Anglosphere destination compared to that in the case of the non-Anglosphere destination. This formalization of friction related to international migration matches the observation that potential migrants prefer destinations which may bring about less anxiety about lower productivity in destination countries due to difficulty in communication stemming from mismatch in language use.

The remainder of this paper is organized as follows. Section 2 presents the model. Section 3 analyzes the stable equilibria and the impact on the stability given by important parameters. Section 4 conducts a welfare analysis, and Section 5 concludes the paper.

2 Model

2.1 Settings

The economy consists of a finite number of equidistant countries $i \in \mathcal{N} = \{1, \cdots, n\}$ $(n \geq 2)$, where countries are divided into two groups, the first of which is labeled the "Anglosphere," and the other "non-Anglosphere," depending on mobility of skilled workers which will be explained in Subsection 2.2. The number of countries in the Anglosphere ($\mathcal{N}_A$) is n_A, and that in the non-Anglosphere ($\mathcal{N}_N$) is n_N, so that the total number of countries satisfies $n = n_A + n_N$. Each country is characterized by its unique (domestic) official language $l \in \mathcal{N}$, and the official language in country i is assumed language i. There are two sectors in the economy, manufacturing (M) and agriculture (A). The manufacturing sector produces horizontally differentiated good, which is transportable across countries under symmetric iceberg trade costs τ, assumed to be symmetric,

$$\tau_{ij} = \begin{cases} 1 & \text{if } i = j \\ \tau > 1 & \text{if } i \neq j. \end{cases}$$

The agriculture sector produces homogeneous good, which can be freely traded across countries.

For workers, there are two types of skill levels, skilled and unskilled. Skilled workers are imperfectly mobile across countries by incurring frictional costs of skill transferring, expressing that the productivity at a foreign destination country is lower than the productivity they can perform in the origin. We formalize this skill transferability in Subsection 2.2. Each skilled worker is endowed with one unit of labor. He is also attached with his own language type l, which is considered his mother tongue. Skilled workers coming from (born in) country i speak language i as their mother tongues, so the official language of the origin country coincides with the mother tongues of skilled workers born there. The population size of the skilled workers with language type l is denoted by H^l, which is assumed to be the same for all language types, so that $H^l = H$ for $l \in \mathcal{N}$. Then the total population of skilled workers in the whole economy is nH. We denote a population size of skilled workers of language type l residing in country i as H_i^l. Put differently, H_i^l is a skilled worker population whose origin is country l and the destination is country i. Analogously, we define $\lambda_i^l \in [0, 1]$ is a fraction of skilled workers residing in country i among skilled workers with language type l, which leads to $\sum_{i=1}^n \lambda_i^l = 1$. We also denote a portfolio of the distribution of skilled workers of language type l across countries $i \in \mathcal{N}$ as $\Lambda^l = (\lambda_1^l, \cdots, \lambda_n^l)$ for $l \in \mathcal{N}$.

Unskilled workers, who are endowed with one unit of labor and inelastically supply it, are immobile across countries[4] but mobile intersectionally. The population size of unskilled workers residing in country i is L_i, which is assumed to be the same across countries, so that $L_i = L$ for all countries, leading to the total unskilled population equal to nL. Then, the total population in country i is $L + \sum_{l=1}^n \lambda_i^l H$. Hereafter, we assume the non-full-specialization (NFS) condition, $\nu \equiv \frac{L}{H} > \frac{\mu(\sigma-1)\sigma}{(1/n)-\mu(\sigma-1)/\sigma}$, to guarantee that the agricultural good is produced in all countries.

2.2 Skill transference

As have been discussed in Section 1, the essence of difficulty in skill transfer in the context of international migration is the mismatch between the official language (dominantly used) in the destination and the mother tongue assigned to each worker, which is, in many cases, the official language in the origin country. This mismatch in language use induces a drop in skilled productivity at the destination country in comparison with the productivity level achieved in the origin. In order to formalize the

[4]Since this paper sheds light on high-skilled migration, for simplicity, we assume low-skilled workers are immobile. However, it may be necessary to support the assumption of low-skilled immobility and high-skilled mobility. From a viewpoint of (potential) migrants' decision in source countries, Grogger and Hanson (2011) indicate that high-skilled workers are more likely to migrate than low-skilled workers by self-selection in a context of migration to OECD destinations. In addition, Chiquiar and Hanson (2005) exhibit a tendency that Mexican migrants to the US are more likely to be better educated than stayers. By these, from a source country side, low-skilled workers show less mobility than high-skilled workers. Further, from a host country side, immigration policies such as the point-based immigration policy may restrict immigrants to those with appropriate educational and economic backgrounds. Moreover, Mayda (2010) suggests that the strictness of immigration policies in host countries affect international migration flows. From these, mobility of low-skilled workers tends to be lower than that of high-skilled workers.

difficulty in skill transfer, or migration costs capturing lower productivity in the destination country, we follow a form used in Grossman and Rossi-Hansberg (2008) and Ottaviano et al. (2013). Specifically, a skilled worker of language type l endowed with one unit of labor can fully supply his labor endowment in the origin country l, but can provide only a smaller fraction of it in the destination country $i \neq l$.

Let δ_i^l be a fraction of labor a skilled worker whose language type is l (=whose origin is county l) can provide in destination country i. Then, δ_i^l expresses an extent of skill transferability by a skilled worker with language type l residing in country i (skill transferability from country l to country i). In other words, δ_i^l is an inverse migration cost because $1 - \delta_i^l$ fraction of labor disappears. Accordingly, we interpret larger δ_i^l implies less frictional migration in terms of productivity. By this, earnings of a skilled worker with language type l residing in country i is given by $y_i = \delta_i^l w_i$, where w_i is a skilled wage in country i, which is endogenously determined in the model.

To relate skill transferability δ_i^l with the Anglosphere/non-Anglosphere, we specify the magnitude order of δ_i^l as:

$$\delta_i^l = \begin{cases} 1 & \text{if } i = l \\ \delta_A^A & \text{if } i \in \mathcal{N}_A \backslash \{l\}, l \in \mathcal{N}_A \\ \delta_A^N & \text{if } i \in \mathcal{N}_A \backslash \{l\}, l \in \mathcal{N}_N \\ \delta_N^A & \text{if } i \in \mathcal{N}_N \backslash \{l\}, l \in \mathcal{N}_A \\ \delta_N^N & \text{if } i \in \mathcal{N}_N \backslash \{l\}, l \in \mathcal{N}_N, \end{cases} \tag{1}$$

where $1 > \delta_A^A > \delta_A^N > \delta_N^A > \delta_N^N$. The reasoning behind this ordering is as follows. First, if a skilled worker is a stayer (not a migrant), he does not incur migration costs and can fully provide his labor endowment in his origin. Next, migration within the Anglosphere is the least frictional among migrations, because migrants from the Anglosphere may suffer from little difficulty in communication in the destination country belonging to the Anglosphere due to the common mother tongue, English, so that migration within the Anglosphere is the least frictional ($1 > \delta_A^A$). The second least frictional migration is the one from the non-Anglosphere to the Anglosphere, because migrants from the non-Anglosphere, having acquired English, may not suffer much from difficulty in communication in the Anglosphere destination but are less fluent than English speakers as a mother tongue ($\delta_A^A > \delta_A^N$). The third least (and better than the worst) frictional migration may be the one from the Anglosphere to the non-Anglosphere. In this case, migrants, who are fluent in English, may suffer from communication difficulty in the non-Anglosphere destination, where a small number of residents can speak English fluently ($\delta_A^N > \delta_N^A$), but this communication difficulty is considered to be less severe than the case of the migration within the non-Anglosphere, where neither migrants nor indigenous residents are fluent in English communication ($\delta_N^A > \delta_N^N$).

2.3 Consumer

By following an n-country footloose entrepreneur model with the quasi-linear upper-tier utility function proposed by Gaspar et al. (2017), the utility function of an individual in country i is given by

$$U_i = \mu \ln M_i + A_i, \qquad \mu > 0, \tag{2}$$

where M_i is the consumption of a CES composite of differentiated varieties of manufactured goods,

$$M_i \equiv \left[\sum_{j=1}^{n} \int_0^{S_j} m_{ji}(s)^{\frac{\sigma-1}{\sigma}} ds \right]^{\frac{\sigma}{\sigma-1}},$$

$m_{ji}(s)$ is the consumption of variety s, produced in country j and consumed in country i, S_j is the number of existing varieties in country j, $\sigma > 1$ is a constant elasticity of substitution between varieties, and A_i is the consumption of the homogeneous agricultural good.

An individual in country i maximizes utility function (2) subject to the following budget constraint

$$\sum_{j=1}^{n} \int_0^{S_j} p_{ji}(s) m_{ji}(s) ds + p_i^A A_i = y_i,$$

where $p_{ji}(s)$ is a delivered price of variety s in country i and produced in country j, and p_i^A is an agricultural good price in country i. The earnings of an individual in country i, y_i, depends on his skill level (skilled or unskilled) and language type, so that $y_i = \delta_i^l w_i$ for skilled workers with language type l residing in country i and $y_i = w_i^L$ for unskilled workers residing in country i, where w_i (w_i^L, resp.) is skilled (unskilled, resp.) wage rate in country i. Utility maximization yields the following demand functions:

$$m_{ji}(s) = \mu \frac{p_{ji}(s)^{-\sigma}}{P_i^{1-\sigma}}, \qquad M_i = \mu P_i^{-1}, \qquad A_i = y_i - \mu, \tag{3}$$

accompanied by the price index of country i associated with the composite good M,

$$P_i \equiv \left[\sum_{k=1}^{n} \int_0^{S_k} p_{ki}(s)^{1-\sigma} ds \right]^{\frac{1}{1-\sigma}}. \tag{4}$$

By substituting (3) to (2), we obtain the indirect utility function (of a skilled worker with language type l) in country i as

$$V_i^{(l)} = y_i^{(l)} - \mu \ln P_i + \mu(\ln \mu - 1), \tag{5}$$

which depends on his language type and where he lives (namely, depends on his origin and his destination).

2.4 Production

Production of agricultural good is inhabited by constant returns to scale and perfect competition. To produce one unit of agricultural good, one unit of unskilled labor is required, and thus, $p_i^A = w_i^L$. Since agricultural good is assumed to be freely traded across countries, it is chosen as the numeraire, which yields $p_i^A = 1$ for $i \in \mathcal{N}$. Then, the unskilled wage is determined as $w_i^L = 1$ for $i \in \mathcal{N}$. By contrast, production of a variety of the manufactured good is inhabited by increasing returns to scale. Production of x units of a variety of the manufactured good is assumed to require input of α units of skilled labor and βx units of unskilled labor, following Forslid and Ottaviano (2003). This leads to the profit function of a firm producing variety s in country i,

$$\pi_i(s) = \sum_{j=1}^{n} m_{ij}(s) \left(L + \sum_{l=1}^{n} \lambda_j^l H \right) [p_{ij}(s) - \tau_{ij}\beta] - \alpha w_i. \tag{6}$$

Profit maximization of (6) with respect to prices yields

$$p_{ii}(s) = \beta \frac{\sigma}{\sigma - 1}, \qquad p_{ij}(s) = \tau\beta \frac{\sigma}{\sigma - 1}. \tag{7}$$

In addition, from skilled labor market clearing for each country, the number of varieties manufactured in country i is

$$S_i = \frac{H}{\alpha} \sum_{l=1}^{n} \delta_i^l \lambda_i^l. \tag{8}$$

By using (7) and (8), the price index (4) is rewritten as

$$P_i = \beta \frac{\sigma}{\sigma - 1} \left(\frac{H}{\alpha} \sum_{j=1}^{n} \phi_{ij} \sum_{l=1}^{n} \delta_j^l \lambda_j^l \right)^{\frac{1}{1-\sigma}}, \tag{9}$$

where freeness of trade between countries i and j is defined as $\phi_{ij} \equiv \tau_{ij}^{1-\sigma} \in (0, 1)$, so that $\phi_{ij} = 1$ if $i = j$ and $\phi_{ij} = \phi$ otherwise.

2.5 Instantaneous equilibrium

Due to free entry and exit, a firm's profit should be zero ($\pi_i = 0$) in the instantaneous equilibrium, so that the skilled wage is determined as

$$w_i = \frac{\mu}{\sigma} \sum_{j=1}^{n} \frac{\phi_{ij} \left(\nu + \sum_{l=1}^{n} \lambda_j^l \right)}{\sum_{k=1}^{n} \phi_{jk} \sum_{l=1}^{n} \delta_k^l \lambda_k^l}. \tag{10}$$

By substituting (8), (9), and (10) into (5), we obtain the indirect utility of a skilled worker with language type l who resides in country i as

$$V_i^l = \frac{\mu}{\sigma} \delta_i^l \sum_{j=1}^{n} \frac{\phi_{ij} \left(\nu + \sum_{m=1}^{n} \lambda_j^m \right)}{\sum_{k=1}^{n} \phi_{jk} \sum_{m=1}^{n} \delta_k^m \lambda_k^m} + \frac{\mu}{\sigma - 1} \ln \left(\sum_{j=1}^{n} \phi_{ij} \sum_{m=1}^{n} \delta_j^m \lambda_j^m \right) + \eta, \tag{11}$$

where a constant $\eta \equiv -\mu \ln \left(\beta \frac{\sigma}{\sigma - 1} \right) + \frac{\mu}{\sigma - 1} \ln H + \mu (\ln \mu - 1)$.[5]

3 Long-run equilibrium

In the long run, skilled workers are mobile across countries and move to a country where he can enjoy a higher indirect utility. In a spatial equilibrium, skilled workers with the same language type l must reach the same utility level V^{l*}:

$$\begin{aligned} V_i^l &= V^{l*} \quad \text{if} \quad \lambda_i^{l*} > 0, \\ V_i^l &\leq V^{l*} \quad \text{if} \quad \lambda_i^{l*} = 0. \end{aligned}$$

[5]Wage is usually considered as one of the fundamentals which affect migration decision. Indirect utility (11) is calculated by using the instantaneous skilled wage (10), so that in our model, wage is treated as an important element of migration decision via the indirect utility.

Candidates of the spatial equilibrium considered are: (i) full agglomeration in the Anglosphere, (ii) full agglomeration in the non-Anglosphere, (iii) dispersion, and (iv) partial agglomeration. For analytical tractability, we focus on the full agglomeration and dispersion. To specify a skilled distribution on the full agglomeration in one country in the Anglosphere or in the non-Anglosphere, we denote country i_C as the industrial core country which attracts all skilled workers existing in the economy. Then, a skilled distribution under the full industrial agglomeration in one country of the Anglosphere, Λ_{aggA}, is expressed by $\Lambda_{aggA} = (\Lambda_1, \cdots, \Lambda_n)$ such that for $l \in \mathcal{N}, i_C \in \mathcal{N}_A$,

$$\lambda_i^l = \begin{cases} 1 & \text{for } i = i_C \\ 0 & \text{for } i \neq i_C. \end{cases} \tag{12}$$

Similarly, a skilled distribution under the full industrial agglomeration in one country of the non-Anglosphere, Λ_{aggN}, is expressed as $\Lambda_{aggN} = (\Lambda_1, \cdots, \Lambda_n)$ such that for $l \in \mathcal{N}, i_C \in \mathcal{N}_N$,

$$\lambda_i^l = \begin{cases} 1 & \text{for } i = i_C \\ 0 & \text{for } i \neq i_C. \end{cases} \tag{13}$$

For a dispersed equilibrium, Λ_{disp}, we define a skilled distribution, $\Lambda_{disp} = (\Lambda^1, \cdots, \Lambda^n)$ such that for $l \in \mathcal{N}$,

$$\lambda_i^l = \begin{cases} 1 & \text{for } i = l \\ 0 & \text{for } i \neq l. \end{cases} \tag{14}$$

In Λ_{disp}, all skilled workers are domestic and each country accommodates the same population of skilled workers, so that skilled workers are equally distributed across countries. That is why Λ_{disp} captures one of the dispersed equilibrium. Below we analyze the stability of skilled distributions, Λ_{aggA}, Λ_{aggN}, and Λ_{disp}.[6]

[6] To complete depicting the stability of the long-run equilibrium, we need to clarify the dynamic system for stability analysis, which expresses the skilled motion in which they move to a country, seeking for the higher indirect utility, such as the replicator dynamics, $\dot{\lambda}_i^l \equiv \frac{d\lambda_i^l}{dt} = \lambda_i^l(V_i^l - V^{l*})$, or the myopic evolutionary dynamics, $\dot{\lambda}_i^l \equiv \frac{d\lambda_i^l}{dt} = \kappa(V_i^l - V^{l*})$, where κ is a positive constant, with $n \times (n-1)$ dynamic equations with $n \times (n-1)$ variables. However, our analysis in this paper is limited to the stability analysis of the equilibria, Λ_{aggA}, Λ_{aggN}, and Λ_{disp}, all of which need only sustain condition analysis for each language type l. Then, we do not determine a specific dynamic system, but only assume that skilled workers move to a country where he can obtain higher utilities.

3.1 Full agglomeration

The aim of this subsection is twofold: (i) we show that Λ_{aggA} is more likely to be stable than Λ_{aggN} if δ_N^N or δ_N^A is small, meaning that the Anglosphere is more likely to attract migrants than the non-Anglosphere if migration within or to the non-Anglosphere is highly frictional; and then (ii) we investigate the stability condition of Λ_{aggA} in terms of important parameters.

The condition for Λ_{aggA} to be stable (the sustain condition for an industrial core in the Anglosphere, $i_C \in \mathcal{N}_A$, to be in stable equilibrium) is to satisfy for all $l \in \mathcal{N}$, $i \in \mathcal{N}\backslash\{i_C\}$

$$V_{i_C}^l|_{\Lambda=\Lambda_{aggA}} > V_i^l|_{\Lambda=\Lambda_{aggA}}, \tag{15}$$

which is shown to reduce to the following condition:[7] for $l \in \mathcal{N}_N$

$$V_{i_C}^l|_{\Lambda=\Lambda_{aggA}} - V_l^l|_{\Lambda=\Lambda_{aggA}} > 0. \tag{16}$$

The reduced sustain condition for Λ_{aggA} (16) indicates that even skilled workers from the non-Anglosphere choose to migrate in the industrial core located in the Anglosphere instead of staying in his home country. This reduction of the sustain condition is natural because (16) comes from the migration decision of skilled workers from the non-Anglosphere, which is the strictest condition to satisfy among migration decision of skilled workers of all language types $l \in \mathcal{N}$ due to the skill transferring difficulty ordering (1). By rewriting the condition (16), we obtain the sustain function f_A as

$$f_A \equiv -\left\{\frac{(n+\nu)\phi^2 + [(n-2)\nu - \delta_A^N n(1+\nu)]\phi + \nu}{\sigma[1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N]\phi} + \frac{\ln\phi}{\sigma-1}\right\}. \tag{17}$$

If $f_A > 0$, Λ_{aggA} is a stable equilibrium. Otherwise, Λ_{aggA} is unstable.

In the same fashion, the sustain condition for Λ_{aggN} to be stable (the sustain condition for an industrial core in the non-Anglosphere, $i_C \in \mathcal{N}_N$, to be in stable equilibrium) is to satisfy for all $l \in \mathcal{N}$, $i \in \mathcal{N}\backslash\{i_C\}$

$$V_{i_C}^l|_{\Lambda=\Lambda_{aggN}} > V_i^l|_{\Lambda=\Lambda_{aggN}}, \tag{18}$$

which is reduced to the condition: for $l \in \mathcal{N}_N$

$$V_{i_C}^l|_{\Lambda=\Lambda_{aggN}} - V_l^l|_{\Lambda=\Lambda_{aggN}} > 0. \tag{19}$$

[7]See Appendix C

Rewriting (19) yields the sustain condition for Λ_{aggN} to be stable associated with the sustain function f_N, where

$$f_N \equiv -\left\{\frac{(n+\nu)\phi^2+[(n-2)\nu-\delta_N^N n(1+\nu)]\phi+\nu}{\sigma[1+(n_N-1)\delta_N^N+(n-n_N)\delta_N^A]\phi}+\frac{\ln\phi}{\sigma-1}\right\}. \tag{20}$$

If $f_N>0$, Λ_{aggN} is stable. Otherwise, Λ_{aggN} is unstable.

3.2 Comparison between the sustain conditions for the full agglomeration in the Anglosphere and non-Anglosphere

By comparing the sustain functions, it is shown that $f_A > f_N$ holds if δ_N^N or δ_N^A is small. More specifically, if $\delta_N^N > \delta^* \equiv \frac{(n-2)\nu}{n(1+\nu)}$ and $\delta_N^A \geq \delta^{**} \equiv \frac{(n-2)\nu+2\sqrt{(n+\nu)\nu}}{n(1+\nu)}$, $f_A - f_N$ is not necessarily positive. Otherwise, $f_A - f_N > 0$ always holds, which is reported in the following proposition.[8]

Proposition 3.1. *If migrating to and within the non-Anglosphere is sufficiently free ($\delta_N^N > \delta^*$ and $\delta_N^A \geq \delta^{**}$), a fully agglomerated equilibrium in the Anglosphere is not necessarily as likely to be stable as a full agglomeration in the non-Anglosphere. Otherwise ($\delta_N^N \leq \delta^*$ or $\delta_N^A < \delta^{**}$), a fully agglomerated equilibrium in the Anglosphere is always more likely to be stable than a full agglomeration in the non-Anglosphere.*

Proposition 3.1 suggests that if migration to the non-Anglosphere were sufficiently free ($\delta_N^N > \delta^*$ and $\delta_N^A \geq \delta^{**}$), the fully agglomerated equilibrium in the non-Anglosphere could be more likely to be stable than the full agglomeration in the Anglosphere. In reality, however, the non-Anglosphere destinations are considered highly frictional for migrants, as have been mentioned in Section 1, this scenario is unlikely. Rather, the second scenario coordinates with the reality of migrants' anxiety about struggling in communication in destination countries. In this scenario, if migration to the non-Anglosphere destination is sufficiently frictional (migrating from the Anglosphere to the non-Anglosphere or migrating within the non-Anglosphere is sufficiently frictional), then a full agglomeration of skilled workers in the Anglosphere is always more likely to be sustained than that in the non-Anglosphere. This finding matches our expectation that less frictional countries in terms of skill transferring for potential migrants (the Anglosphere) can attract more international migrants and more likely to become industrial core than more frictional countries (the non-Anglosphere). This follows the empirical assertion that English-speaking countries can attract more international migrants than non-English-speaking

[8]See Appendix D for the proof.

countries because skill transfer is smoother when English-speaking countries are chosen as the destination

3.3 Full agglomeration in the Anglosphere and important parameters

We inspect the impacts of important parameters, such as $\delta_A^N, \phi, \delta_A^A$, and n_A, on the the stability of the fully agglomerated equilibrium in the Anglosphere. As has been mentioned in Section 1, there are two options in how to interpret "Anglosphere." In the normal interpretation, the Anglosphere consists of English-speaking countries where English is adopted as an official language, but in the extended interpretation, the Anglosphere consists of countries where English communication is not as difficult, such as European countries as well as English-speaking countries. With the adoption of the extended interpretation of the Anglosphere, δ_A^A and n_A can change as well as δ_A^N, because spread of English communication comes from (i) expansion of the Anglosphere (if the number of countries within the Anglosphere increases then English communication is smooth in more countries), captured by an increase in n_A, (ii) less frictional migration from the non-Anglosphere to the Anglosphere, captured by an increase in δ_A^N, and (iii) less frictional migration within the Anglosphere, captured by an increase in δ_A^A. Following the extended interpretation of the Anglosphere, we first investigate the impacts of the spread of English on the sustain condition (stability condition of an equilibrium of industrially agglomeration in the Anglosphere), and then move to a discussion related to an impact of trade freeness, ϕ, on the sustain function f_A.

For the impact given by δ_A^N, because $\frac{\partial f_A}{\partial \delta_A^N} > 0$,[9] it immediately yields the following proposition.

Proposition 3.2. *If migration from the non-Anglosphere becomes less frictional (larger δ_A^N), then the agglomerated equilibrium in an industrial core country of the Anglosphere is more likely to be stable (sustain condition for Λ_{aggA}, $f_A > 0$, is more likely to hold).*

To interpret Proposition 3.2 and seek for a mechanism behind it, we consider two opposite forces driven by an increase in δ_A^N, one of which is related to market crowding effect (competition among firms) and the other is related to less frictional migration.

First, we consider the impact of an increase in δ_A^N on f_A in relation with the market crowding effect. When δ_A^N increases, then the market crowding effect (competition among firms) in country i_C becomes severe (this increases the denominator of the first term in f_A). Because this strengthens dispersion force in the economy and firms (skilled workers) prefer to circumvent severe competition,

[9]See Appendix E

this goes in a direction in which it is less likely for the industrially agglomerated equilibrium in the Anglosphere (Λ_{aggA}) to be stable. Next, from a viewpoint of less frictional migration, we investigate the impact given by a rise in δ_A^N. Since an increase in δ_A^N indicates that migrating from the non-Anglosphere to the Anglosphere becomes less frictional, migrants can enjoy benefits of agglomeration (market size effect) without suffering from a severe decline in income caused by lower productivity in the destination country, which leads to a situation in which the full agglomeration in the Anglosphere (Λ_{aggA}) is more likely to be stable. By comparing these two impacts of an increase in δ_A^N, the finding that $\frac{\partial f_A}{\partial \delta_A^N} > 0$ always holds comes from that the positive impact of the market size effect under less frictional migration always exceeds the negative impact of the market crowding effect. Then, an increase in δ_A^N enhances the stability of a full agglomeration in the Anglosphere.

For the impact given by a change in δ_A^A, we obtain the following proposition, by inspecting $\frac{\partial f_A}{\partial \delta_A^A}$.[10]

Proposition 3.3. *If possibility of inducing the market crowding effect is severe ($\delta_A^N > \delta^*$ and $\delta_A^A \geq \delta^{**}$), less frictional migration within the Anglosphere does not necessarily stabilize the fully agglomerated equilibrium in the Anglosphere. Otherwise (possibility of inducing the market crowding is not severe; $\delta_A^N \leq \delta^*$ or $\delta_A^A < \delta^{**}$), less frictional migration within the Anglosphere always stabilizes the fully agglomerated equilibrium in the Anglosphere.*

A mechanism behind Proposition 3.3 can be grasped by the comparison between the market crowding effect and the market size effect, rather than the originally motivated interpretation of skill transferring friction. Remember that the sustain condition $f_A > 0$ comes from migration decision of skilled workers with the non-Anglosphere origin, but not the decision of skilled workers with the Anglosphere origin. Then, a change in δ_A^A does not give an impact on a change in f_A through the channel of "less frictional migration effect" unlike in the case of a change in δ_A^N. By increasing δ_A^A, the market crowding effect (a dispersion force) as well as the market size effect (an agglomeration force) becomes stronger. Given these positive and negative impacts of an increase in δ_A^A on the stability of agglomeration, consider what $\delta_A^N > \delta^*$ and $\delta_A^A \geq \delta^{**}$ mean. Because large δ_A^N and δ_A^A imply higher possibility of inducing the severe market crowding effect by increasing the denominator of f_A, when both δ_A^N and δ_A^A are large ($\delta_A^N > \delta^*$ and $\delta_A^A \geq \delta^{**}$), the fully agglomerated equilibrium is not necessarily stabilized even when δ_A^A increases. By contrast, if there is little possibility of suffering from market crowding, an increase in δ_A^A stabilizes the skilled agglomeration in the Anglosphere ($f_A > 0$ more likely to be satisfied) through a channel of enjoying market size effect because the agglomeration force

[10] See Appendix E.

stemming from the market size effect exceeds the dispersion force borne by the market crowding effect.

A discussion on the impact of an increase in n_A on f_A is almost the same as in the case of δ_A^A. By inspecting $\frac{\partial f_A}{\partial n_A}$, we obtain the following proposition.[11]

Proposition 3.4. *If possibility of inducing the market crowding effect is severe ($\delta_A^N > \delta^*$ and $\delta_A^A \geq \delta^{**}$), an expansion of the Anglosphere does not necessarily stabilize the fully agglomerated equilibrium in the Anglosphere. Otherwise (possibility of inducing the market crowding is not severe; $\delta_A^N \leq \delta^*$ or $\delta_A^A < \delta^{**}$), an expansion of the Anglosphere always stabilizes the fully agglomerated equilibrium in the Anglosphere.*

Proposition 3.4 has the same implication as that of Proposition 3.3, because an increase in n_A has the same impact as an increasing in δ_A^A in that it has an agglomeration force through the market size effect as well as the dispersion force through the market crowding effect.

Summing up the impacts of parameters related to the spread of English communication, slightly different tendencies are found in the case of a change in δ_A^N, which is related to the spread of English for the non-Anglosphere countries, and the case of a change in δ_A^A and n_A, which are related to the spread of English for the Anglosphere countries. For skilled workers whose origin country is in the non-Anglosphere, less frictional migration to the Anglosphere in terms of skill transference always accelerates industrial agglomeration in the Anglosphere. On the other hand, less frictional migration within the Anglosphere and expansion of the Anglosphere do not necessarily accelerate industrial agglomeration in the Anglosphere.

Finally, we look at a relationship between the sustain function f_A and trade freeness ϕ. Because it is shown that $f_A|_{\phi\to 0} < 0$, $f_A|_{\phi\to 1} < 0$, and f_A has a unique maximum at $\phi = \phi^*$ in the interval $\phi \in (0,1)$ such that $\frac{\partial f_A}{\partial \phi} \geq 0$ for $\phi \leq \phi^*$ and $\frac{\partial f_A}{\partial \phi} < 0$ for $\phi > \phi^*$, we obtain the following proposition in which we define the maximum value of f_A in $\phi \in (0,1)$ as $f_A^{\max} \equiv f_A|_{\phi=\phi^*}$.[12]

Proposition 3.5. *When $f_A^{\max} \leq 0$, agglomeration in the Anglosphere is unstable regardless of ϕ ($f_A \leq 0$ at any level of ϕ). When $f_A^{\max} > 0$, agglomeration in the Anglosphere is stable for the intermediate level of trade freeness ($f_A > 0$ for $\phi \in (\underline{\phi}, \bar{\phi})$), while agglomeration in the Anglosphere is unstable at high and low levels of trade freeness ($f_A \leq 0$ for $\phi \notin (\underline{\phi}, \bar{\phi})$).*

Before interpreting Proposition 3.5, recall that an increase in f_A is most importantly caused by an increase in δ_A^N, as asserted in Proposition 3.2, because an increase in δ_A^N always shifts f_A upwards.

[11]See Appendix E.
[12]See Appendix E.

This leads to an observation that the case of $f_A^{\max} > 0$ will basically correspond to a situation of high δ_A^N (less frictional for migrants from the non-Anglosphere), and the case of $f_A^{\max} \leq 0$ a situation of low δ_A^N (more frictional for migrants from the non-Anglosphere).

When δ_A^N is low so that migration from the non-Anglosphere to the Anglosphere is highly frictional ($f_A^{\max} \leq 0$), Λ_{aggA} cannot be stable at any level of trade freeness. On the other hand, when δ_A^N is sufficiently high so that migration from the non-Anglosphere to the Anglosphere is not severely frictional ($f_A^{\max} > 0$), Λ_{aggA} is stable for an intermediate level of trade freeness ($\phi \in (\underline{\phi}, \bar{\phi})$) but is unstable for high and low levels of trade freeness ($\phi \notin (\underline{\phi}, \bar{\phi})$). In this case, there are two transitions of a sign of f_A, the first of which is around $\phi = \underline{\phi}$ and the second around $\phi = \bar{\phi}$. The first transition around $\phi = \underline{\phi}$ (switch from a negative to positive sign of f_A) is caused by a decline in the trade cost τ. Because a decrease in τ leads shipping not too costly, the industrial agglomeration more likely to be in stable equilibrium. The second transition around $\phi = \bar{\phi}$ (switch from a positive to negative sign of f_A) is caused by the existence of δ_i^l. Without $\delta_i^l < 1$ (i.e., $\delta_i^l = 1$, especially $\delta_A^N = 1$) unlike in our model, completely free trade ($\phi = 1$) should lead to $V_i^l = V_j^l$ for all i and j. However, skill transferring cost, $1 - \delta_i^l$, still exists and works as a dispersion force even under highly free trade circumstances in the face of $\delta_i^l < 1$ in the present model. This leads to the result that the full agglomeration is unstable even when trade freeness is high ($\phi > \bar{\phi}$) when international migration is inhabited by skill transferring friction.

3.4 Dispersion

For the final possibility of the long-run equilibrium at hand, we investigate stability of the dispersed equilibrium. The condition for Λ_{disp} to be stable is

$$V_l^l|_{\Lambda=\Lambda_{disp}} > V_i^l|_{\Lambda=\Lambda_{disp}} \qquad \text{for } l \in \mathcal{N}, i \in \mathcal{N}\backslash\{l\}. \tag{21}$$

Because $\Delta V^l|_{\Lambda=\Lambda_{disp}} \equiv V_l^l|_{\Lambda=\Lambda_{disp}} - V_i^l|_{\Lambda=\Lambda_{disp}} = \frac{\mu}{\sigma}(1-\delta_i^l)(1+\nu) > 0$, the stability condition of Λ_{disp} is always satisfied, which yields the following proposition.

Proposition 3.6. *With frictional migration costs which lowers productivity at the destination, dispersed configuration, Λ_{disp}, is always in stable equilibrium.*

Because dispersion always consists of a stable equilibrium, a situation in which all skilled workers are domestic workers and international migration does not take place can stably occur.

4 Social welfare

Since there is possibility of multiple equilibria as discussed in Section 3, we analyze the desirability of these equilibria based on the utilitarian criterion and compare the average indirect utilities for each equilibrium. We consider the desirability by group (skilled and unskilled workers) and that for the society as a whole.

4.1 Skilled workers

To compare the desirability of long-run equilibria, $\Lambda_{aggA}, \Lambda_{aggN}$, and Λ_{disp}, for skilled workers, we use the average indirect utility of skilled workers,

$$\bar{V} \equiv \frac{1}{n}\sum_{l=1}^{n}\bar{V}^{l} \equiv \frac{1}{n}\sum_{l=1}^{n}\sum_{i=1}^{n}\lambda_i^l V_i^l, \tag{22}$$

and discuss which equilibrium is the most desirable for skilled workers by comparing $\bar{V}|_{\Lambda=\Lambda_{aggA}}$, $\bar{V}|_{\Lambda=\Lambda_{aggN}}$, and $\bar{V}|_{\Lambda=\Lambda_{disp}}$. First, we make a comparison between Λ_{aggA} and Λ_{aggN}, which reveals that $\bar{V}|_{\Lambda=\Lambda_{aggA}} > \bar{V}|_{\Lambda=\Lambda_{aggN}}$ always holds.[13] Then, we obtain the following lemma.

Lemma 4.1. *Agglomeration in the Anglosphere is always more desirable than agglomeration in the non-Anglosphere for skilled workers.*

From Lemma 4.1, the comparison reduces to that between Λ_{aggA} and Λ_{disp}. By comparing Λ_{aggA} and Λ_{disp},

$$\bar{V}|_{\Lambda=\Lambda_{aggA}} \gtreqless \bar{V}|_{\Lambda=\Lambda_{disp}} \Leftrightarrow g^H \gtreqless \phi, \tag{23}$$

where $g^H \equiv [(n_A-1)\delta_A^A + (n-n_A)\delta_A^N]/(n-1)$, which can be interpreted as the weighted average of freeness of migration to/within the Anglosphere. By inspecting (23), we obtain the following proposition.

Proposition 4.1. *If migration to/within the Anglosphere is relatively freer than trade ($g^H > \phi$), then agglomeration in the Anglosphere is more desirable than dispersion for skilled workers ($\bar{V}|_{\Lambda=\Lambda_{aggA}} > \bar{V}|_{\Lambda=\Lambda_{disp}}$). If trade is relatively freer than migration to/within the Anglosphere ($g^H < \phi$), then dispersion is more desirable than agglomeration in the Anglosphere for skilled workers ($\bar{V}|_{\Lambda=\Lambda_{aggA}} < \bar{V}|_{\Lambda=\Lambda_{disp}}$).*

[13] See Appendix F.

In addition, $\frac{\partial g^H}{\partial \delta_A^A} > 0$, $\frac{\partial g^H}{\partial \delta_A^N} > 0$, and $\frac{\partial g^H}{\partial n_A} > 0$ always hold, which yields that if migration to the Anglosphere is less frictional, or the Anglosphere expands, then it is more likely that agglomeration in the Anglosphere is more desirable than dispersion for skilled workers.

4.2 Unskilled workers

By taking the similar procedure in Subsection 4.1, we calculate the average indirect utility of unskilled workers,

$$\bar{V}^L \equiv \frac{1}{n} \sum_{i=1}^{n} V_i^L, \tag{24}$$

where V_i^L is an indirect utility of an unskilled worker residing in country i. By using (24), we compare the desirability of equilibria, $\Lambda_{aggA}, \Lambda_{aggN}$, and Λ_{disp}, for unskilled workers. Comparing Λ_{aggA} and Λ_{aggN} verifies that $\bar{V}|_{\Lambda=\Lambda_{aggA}} > \bar{V}|_{\Lambda=\Lambda_{aggN}}$ always holds, which leads to the following lemma as in the case of skilled workers.[14]

Lemma 4.2. *Agglomeration in the Anglosphere is always more desirable than agglomeration in the non-Anglosphere for unskilled workers.*

Lemma 4.2 reduces the comparison among $\Lambda_{aggA}, \Lambda_{aggN}$, and Λ_{disp} to that between Λ_{aggA} and Λ_{disp}. By comparing Λ_{aggA} and Λ_{disp}, it is shown that $\bar{V}^L|_{\Lambda=\Lambda_{disp}} > \bar{V}^L|_{\Lambda=\Lambda_{aggA}}$ always holds.[15] Then we obtain the following proposition.[16]

Proposition 4.2. *For unskilled workers, dispersion is always more desirable than agglomeration in the Anglosphere (and hence, than agglomeration in the non-Anglosphere).*

4.3 Society

Finally ,we analyze what distribution, $\Lambda_{aggA}, \Lambda_{aggN}$, or Λ_{disp}, is the most desirable for the whole economy. The social welfare function, the average indirect utility of all individuals in the economy, is

$$W \equiv \frac{1}{1+\nu}\bar{V} + \frac{\nu}{1+\nu}\bar{V}^L. \tag{25}$$

From Lemmas 4.1 and 4.2, and using that W is a linear combination of $\bar{V}$ and $\bar{V}^L$, we immediately obtain the following lemma.

[14] See Appendix F.
[15] Accordingly, for unskilled workers remaining in a peripheral country $i \neq i_C$, $V_i^L|_{\Lambda=\Lambda_{disp}} > V_i^L|_{\Lambda=\Lambda_{aggA}}$.
[16] See Appendix F.

Lemma 4.3. *Agglomeration in the Anglosphere is always socially more desirable than agglomeration in the non-Anglosphere in view of social welfare.*

Here again, the comparison reduces to that between Λ_{aggA} and Λ_{disp}, and we investigate which distribution is socially more desirable. From the comparison between Λ_{aggA} and Λ_{disp},

$$W|_{\Lambda=\Lambda_{aggA}} \gtreqless W|_{\Lambda=\Lambda_{disp}} \Leftrightarrow g \gtreqless 0, \tag{26}$$

where $g \equiv \left\{ n(1+\nu)\ln\left[\frac{1+(n_A-1)\delta_A^A+(n-n_A)\delta_A^N}{1+(n-1)\phi}\right] + (n-1)\nu\ln\phi \right\}$. Below, by inspecting g in terms of important parameters, $\delta_A^N, \delta_A^A, n_A$, and ϕ, we analyze which of Λ_{aggA} and Λ_{disp} is more likely to be socially desirable.

As for parameters related to the Anglosphere, because $\frac{\partial g}{\partial \delta_A^A} > 0$, $\frac{\partial g}{\partial \delta_A^N} > 0$, and $\frac{\partial g}{\partial n_A} > 0$ always hold,[17] it can be asserted that freer migration to the Anglosphere, freer migration within the Anglosphere, and expansion of the Anglosphere lead to full agglomeration in the Anglosphere more likely to be socially desirable than dispersion.

Turning to trade freeness ϕ and g, it is shown that $g|_{\phi\to 0} < 0$, $g|_{\phi\to 1} < 0$, and g has a unique maximum, $g^{\max} \equiv g|_{\phi=\phi^{W*}}$, at $\phi = \phi^{W*} \equiv \frac{\nu}{n+\nu}$ such that $\frac{\partial g}{\partial \phi} > 0$ for $\phi < \phi^{W*}$ and $\frac{\partial g}{\partial \phi} \leq 0$ for $\phi \geq \phi^{W*}$, which yields the following proposition[18].

Proposition 4.3. *When $g^{\max} < 0$, agglomeration in the Anglosphere is socially less desirable than dispersion ($W|_{\Lambda=\Lambda_{aggA}} < W|_{\Lambda=\Lambda_{disp}}$) at any level of ϕ. When $g^{\max} > 0$, agglomeration in the Anglosphere is socially more desirable than dispersion ($W|_{\Lambda=\Lambda_{aggA}} > W|_{\Lambda=\Lambda_{disp}}$) for an intermediate $\phi \in (\underline{\phi}^W, \bar{\phi}^W)$, while agglomeration in the Anglosphere is socially less desirable than dispersion ($W|_{\Lambda=\Lambda_{aggA}} < W|_{\Lambda=\Lambda_{disp}}$) for high and low $\phi \notin [\underline{\phi}^W, \bar{\phi}^W]$*

By taking into account that $\frac{\partial g}{\partial \delta_A^A} > 0$, $\frac{\partial g}{\partial \delta_A^N} > 0$, and $\frac{\partial g}{\partial n_A} > 0$, all of which lead to Λ_{aggA} more likely to be socially superior to Λ_{disp}, Proposition 4.3 can be interpreted in a way that less frictional migration to/within the Anglosphere or an expansion of the Anglosphere leads to a possibility of the full agglomeration in the Anglosphere more desirable than the dispersion at an intermediate level of trade freeness. Conversely, more frictional migration to/within the Anglosphere or the shrinkage of the Anglosphere can lower the possibility that the full agglomeration in the Anglosphere more desirable than the dispersion at any level of trade freeness.

[17]See Appendix F.
[18]See Appendix F.

The final comment in this section is related to the comparison between the equilibrium and the socially desirable distributions. In Subsection 3.3, we find that both Λ_{aggA} and Λ_{disp} are stable for $\phi \in (\underline{\phi}, \bar{\phi})$ (intermediate level of trade freeness) when $f_A^{\max} > 0$ (sufficiently frictionless in skill transferring, especially for migrants from the non-Anglosphere to the Anglosphere). Then, in relation with immigration policies associated with skill transference hindered by language mismatch, if the realized industrial distribution in equilibrium is Λ_{disp} under less severe migration costs for skilled workers from the non-Anglosphere and intermediate trade costs, it may be socially desirable to switch from Λ_{disp} to Λ_{aggA} by providing language training for immigrants from the non-Anglosphere to lessen the language barriers and make them to assimilate easier.

5 Conclusion

This paper analyzed how skill transference in migration affects spatial equilibrium configuration of skilled workers. In international migration, potential migrants may suffer from lower productivity at destination countries due to mismatch in language use. Especially, by spotlighting the empirical findings in the previous literature showing that English-speaking countries tend to be chosen as the destination by international migrants rather than non-English-speaking countries, we constructed a model to explain how frictional skill transfer in migration affects the mobile skilled worker distribution across countries. Based on an n-country footloose entrepreneur model proposed by Gaspar et al. (2017), we introduced an asymmetric skill transference element in the model, in which asymmetry of migration costs stems from to which category of the linguistic areas, the Anglosphere and the non-Anglosphere, the origin and destination countries belong. Specifically, the Anglosphere is assumed to be less frictional countries in terms of skill transference for skilled workers while the non-Anglosphere are more friction-inducing countries.

The main result we derived is that if migration to the non-Anglosphere countries is sufficiently frictional, the fully agglomerated equilibrium in the Anglosphere is more likely to be stable than that in the non-Anglosphere. Also, we showed that the parameter related to migrants from the non-Anglosphere and those related to migrants from the Anglosphere tend to reveal slightly different tendencies on how to affect the stability condition of the full agglomeration in the Anglosphere. For skilled workers whose origin country is in the non-Anglosphere, less frictional migration to the Anglosphere in terms of skill transference always enhances the stability of an equilibrium of the full agglomeration in the Anglosphere. By contrast, less frictional migration within the Anglosphere and the expansion of the Anglosphere (a larger number of countries where English communication is sufficiently smooth) do not necessarily enhance the stability of an equilibrium of the full agglomeration in the Anglosphere. In addition, due to the existence of friction in international skill transference, dispersed equilibrium, in which all skilled workers reside in their home country and work as domestic workers, is always stable.

From a viewpoint of social welfare, in the circumstance under an intermediate level of trade freeness and sufficiently frictionless skill transference, full agglomeration in the Anglosphere is socially superior to dispersion. In equilibrium, on the other hand, both distributions, agglomeration in the Anglosphere and dispersion, can be stable. Then, if the realized skilled distribution is the dispersed stable equilibrium, it may be desirable to implement a policy to switch from dispersion to agglomeration, for instance, by providing language training for immigrants from the non-Anglosphere to lessen the language barriers they face and make them to assimilate easier.

Finally, a future extension of this paper is mentioned. In our present model, decision about how much potential migrants learn English is not considered. However, it may be more appropriate to consider this point because the extent of skill transferring friction caused by language difference is affected by an English acquisition level of migrants. Then, potential migrants may make an effort to acquire English before migration in order to circumvent skill transferring friction at destination. From this viewpoint, English education can be considered as an investment, so that it may be valuable to construct a model in which potential migrants endogenously decide whether to learn English or not, or the effort level they devote for English acquisition before migration. Although this is beyond this paper, it is worth to go in this direction for the future work.

References

Adsera, A. and M. Pytlikova (2015): "The Role of Language in Shaping International Migration," *The Economic Journal*, 125, F49–F81.

Artuc, E., F. Docquier, Ç. Özden, and C. Parsons (2015): "A Global Assessment of Human Capital Mobility: The Role of Non-OECD Destinations," *World Development*, 65, 6–26.

Belot, M. and S. Ederveen (2012): "Cultural Barriers in Migration between OECD Countries," *Journal of Population Economics*, 25, 1077–1105.

Belot, M. V. and T. J. Hatton (2012): "Immigrant Selection in the OECD," *The Scandinavian Journal of Economics*, 114, 1105–1128.

Chiquiar, D. and G. H. Hanson (2005): "International Migration, Self-selection, and the Distribution of Wages: Evidence from Mexico and the United States," *Journal of Political Economy*, 113, 239–281.

Chiswick, B. R. and P. W. Miller (2014): "International Migration and the Economics of Language," *Handbook of the Economics of Immigration*, 211–269.

Clark, X., T. J. Hatton, and J. G. Williamson (2007): "Explaining US Immigration, 1971–1998," *The Review of Economics and Statistics*, 89, 359–373.

Dustmann, C. and A. Van Soest (2001): "Language Fluency and Earnings: Estimation with Misclassified Language Indicators," *Review of Economics and Statistics*, 83, 663–674.

——— (2002): "Language and the Earnings of Immigrants," *Industrial and Labor Relations Review*, 55, 473–492.

Forslid, R. and G. I. Ottaviano (2003): "An Analytically Solvable Core-periphery Model," *Journal of Economic Geography*, 3, 229–240.

Gaspar, J. M., S. B. S. D. Castro, and J. Correia-da Silva (2017): "Agglomeration Patterns In A Multi-regional Economy Without Income Effects," *Economic Theory*.

Grogger, J. and G. H. Hanson (2011): "Income Mmaximization and the Selection and Sorting of International Migrants," *Journal of Development Economics*, 95, 42–57.

Grossman, G. M. and E. Rossi-Hansberg (2008): "Trading Tasks: A Simple Theory of Offshoring," *American Economic Review*, 98, 1978–1997.

Ludema, R. D. and I. Wooton (1999): "Regional Integration, Trade and Migration: Are Demand Linkages Relevant in Europe?" *Migration: The Controversies and the Evidence*, 51–68.

Mayda, A. M. (2010): "International Migration: A Panel Data Analysis of the Determinants of Bilateral Flows," *Journal of Population Economics*, 23, 1249–1274.

Ottaviano, G. I., G. Peri, and G. C. Wright (2013): "Immigration, Offshoring, and American jobs," *American Economic Review*, 103, 1925–1959.

Pedersen, P. J., M. Pytlikova, and N. Smith (2008): "Selection and Network Effects—Migration Flows into OECD Countries 1990–2000," *European Economic Review*, 52, 1160–1186.

Roy, A. D. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3, 135–146.

Tabuchi, T., J.-F. Thisse, and X. Zhu (2017): "Does Technological Progress Magnify Regional Disparities?" *International Economic Review (forthcoming)*.

Appendix A Long-run trend of immigrant inflows to English-speaking/non-English-speaking countries

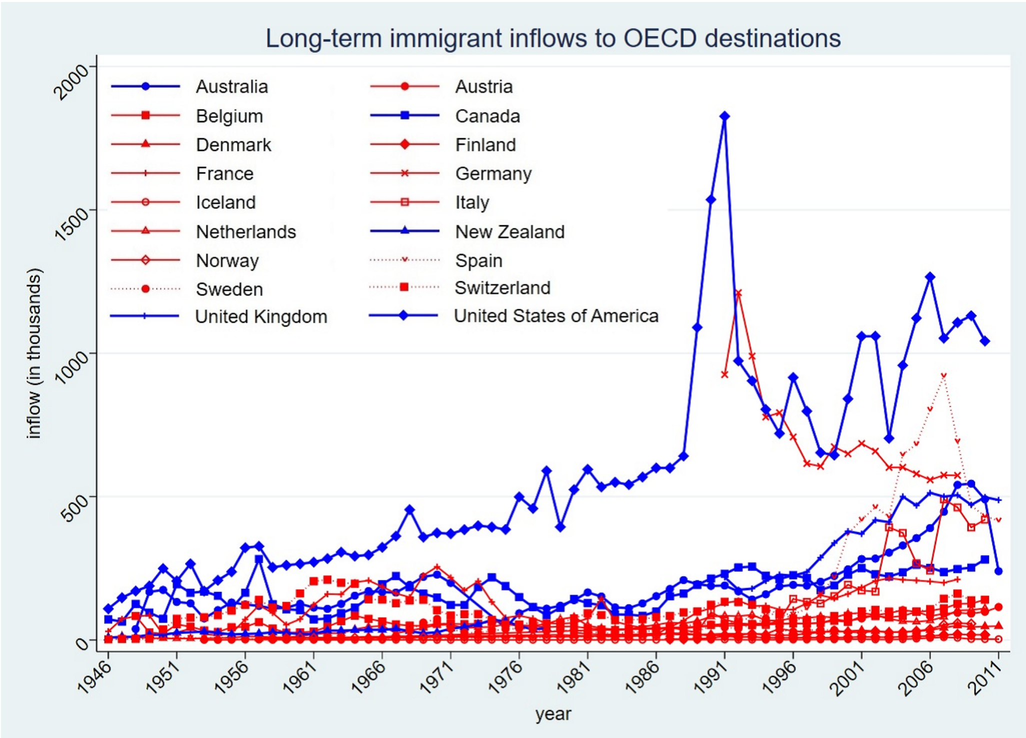

Figure 1: Long-run trend of immigrant inflows to the OECD destinations
DEMIG (2015) DEMIG TOTAL, version 1.5. Oxford: International Migration Institute, University of Oxford. www.migrationdeterminants.eu

Appendix B Skilled migration tendency within/between the Anglosphere and non-Anglosphere

Table 1: Share of skilled migrants to the total migrants to OECD destinations, decomposed according to Anglosphere/non-Anglosphere (%)

Year\Origin-Destination	"Anglo"-"Anglo"	"Non"-"Anglo"	"Anglo"-"Non"	"Non"-"Non"
2000	55.7	41.5	35.7	19.3
1990	44.1	37.6	30.9	14.8

"Anglo"=Anglosphere, "Non"=Non-Anglosphere
Figures in the table are shares of "immigrant stock to OECD destinations with college degree whose age of entry is 22+" to "immigrant stock to OECD destinations whose age of entry is 22+."
Artuc et al. (2015) https://perso.uclouvain.be/frederic.docquier/oxlight.htm

Appendix C Sustain condition for Λ_{aggA}

We show the sustain condition for Λ_{aggA} (15) is reduced to the condition (16) by dividing the discussions into two steps. In the first step, we separately calculate indirect utilities for $l = i_C$, $l \in \mathcal{N}_A \backslash \{i_C\}$, and $l \in \mathcal{N}_N$ under the skilled distribution Λ_{aggA}. Then, for each case, we compare the indirect utilities and obtain a sustain condition, i.e., for language types $l = i_C$, $l \in \mathcal{N}_A \backslash \{i_C\}$, and $l \in \mathcal{N}_N$, we obtain the condition that residing in country i_C bears a higher indirect utility than any other countries. The second step is to find the strictest sustain condition (least likely to take positive values) among language types $l = i_C$, $l \in \mathcal{N}_A \backslash \{i_C\}$, and $l \in \mathcal{N}_N$. Then, extract the strictest case as the final sustain function for Λ_{aggA} to be stable. (When the final sustain condition (the strictest condition) is satisfied, residing in country i_C bears the highest indirect utility for all skilled workers with any language type.) Below, a constant η is omitted unless necessary.

(i) $l = i_C$

For skilled workers with language type $l = i_C$, the indirect utility under Λ_{aggA} is

$$V_i^l|_{\Lambda=\Lambda_{aggA}, l=i_C} = \begin{cases} \dfrac{\mu}{\sigma} \dfrac{n(1+\nu)}{1+(n_A-1)\delta_A^A+(n-n_A)\delta_A^N} + \dfrac{\mu}{\sigma-1} \ln\left[1+(n_A-1)\delta_A^A+(n-n_A)\delta_A^N\right] \\ \hfill (\text{for } i = i_C) \quad (27a) \\ \dfrac{\mu}{\sigma} \dfrac{\delta_A^A[(n+\nu)\phi^2+(n-2)\nu\phi+\nu]}{[1+(n_A-1)\delta_A^A+(n-n_A)\delta_A^N]\phi} + \dfrac{\mu}{\sigma-1} \ln\left\{[1+(n_A-1)\delta_A^A+(n-n_A)\delta_A^N]\phi\right\} \\ \hfill (\text{for } i \in \mathcal{N}_A \backslash \{i_C\}) \quad (27b) \\ \dfrac{\mu}{\sigma} \dfrac{\delta_N^A[(n+\nu)\phi^2+(n-2)\nu\phi+\nu]}{[1+(n_A-1)\delta_A^A+(n-n_A)\delta_A^N]\phi} + \dfrac{\mu}{\sigma-1} \ln\left\{[1+(n_A-1)\delta_A^A+(n-n_A)\delta_A^N]\phi\right\} \\ \hfill (\text{for } i \in \mathcal{N}_N). \quad (27c) \end{cases}$$

The condition on whether residing in country i_C is the best for skilled workers of language type $l = i_C \in \mathcal{N}_A$ is reduced to a comparison between (27a) and (27b) because by $\delta_A^A > \delta_N^A$, (27b) is greater than (27c). Then, the reduced sustain condition for skilled workers with language type $l = i_C$ is

$$\Delta V^{i_C}|_{\Lambda=\Lambda_{aggA}} \equiv V_i^l|_{\Lambda=\Lambda_{aggA}, l=i_C, i=i_C} - V_i^l|_{\Lambda=\Lambda_{aggA}, l=i_C, i\in\mathcal{N}_A\backslash\{i_C\}} > 0. \tag{28}$$

(ii) $l \in \mathcal{N}_A \backslash \{i_C\}$

For skilled workers with language type $l \in \mathcal{N}_A \backslash \{i_C\}$, the indirect utility under Λ_{aggA} is

$$V_i^l|_{\Lambda=\Lambda_{aggA}, l \in \mathcal{N}_A \backslash \{i_C\}} = \begin{cases} \dfrac{\mu}{\sigma} \dfrac{\delta_A^A n(1+\nu)}{1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N} + \dfrac{\mu}{\sigma-1} \ln\left[1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N\right] & \text{(for } i = i_C\text{)} \quad (29a) \\ \dfrac{\mu}{\sigma} \dfrac{\phi^2(n+\nu)+\phi(n-2)\nu+\nu}{[1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N]\phi} + \dfrac{\mu}{\sigma-1} \ln\left\{[1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N]\phi\right\} & \text{(for } i = l\text{)} \quad (29b) \\ \dfrac{\mu}{\sigma} \dfrac{\delta_A^A[(n+\nu)\phi^2+(n-2)\nu\phi+\nu]}{[1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N]\phi} + \dfrac{\mu}{\sigma-1} \ln\left\{[1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N]\phi\right\} & \text{(for } i \in \mathcal{N}_A \backslash \{i_C, l\}\text{)} \quad (29c) \\ \dfrac{\mu}{\sigma} \dfrac{\delta_N^A[(n+\nu)\phi^2+(n-2)\nu\phi+\nu]}{[1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N]\phi} + \dfrac{\mu}{\sigma-1} \ln\left\{[1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N]\phi\right\} & \text{(for } i \in \mathcal{N}_N\text{).} \quad (29d) \end{cases}$$

By $1 > \delta_A^A > \delta_N^A$, (29b)>(29c)>(29d) always holds. Then, the condition on whether residing in country i_C is the best for skilled workers of language type $l \in \mathcal{N}_A \backslash \{i_C\}$ is reduced to a comparison between (29a) and (29b), which leads to that the condition for skilled workers with language type $l \in \mathcal{N}_A \backslash \{i_C\}$ is

$$\Delta V^{\mathcal{N}_A \backslash \{i_C\}}|_{\Lambda=\Lambda_{aggA}} \equiv V_i^l|_{\Lambda=\Lambda_{aggA}, l \in \mathcal{N}_A \backslash \{i_C\}, i=i_C} - V_i^l|_{\Lambda=\Lambda_{aggA}, l \in \mathcal{N}_A \backslash \{i_C\}, i=l} > 0. \quad (30)$$

(iii) $l \in \mathcal{N}_N$

For skilled workers with language type $l \in \mathcal{N}_N$, the indirect utility under Λ_{aggA} is

$$V_i^l|_{\Lambda=\Lambda_{aggA}, l \in \mathcal{N}_N} = \begin{cases} \dfrac{\mu}{\sigma} \dfrac{\delta_A^N n(1+\nu)}{1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N} + \dfrac{\mu}{\sigma-1} \ln\left[1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N\right] & \text{(for } i = i_C\text{)} \quad (31a) \\ \dfrac{\mu}{\sigma} \dfrac{\phi^2(n+\nu)+\phi(n-2)\nu+\nu}{[1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N]\phi} + \dfrac{\mu}{\sigma-1} \ln\left\{([1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N]\phi\right\} & \text{(for } i = l\text{)} \quad (31b) \\ \dfrac{\mu}{\sigma} \dfrac{\delta_A^N[(n+\nu)\phi^2+(n-2)\nu\phi+\nu]}{[1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N]\phi} + \dfrac{\mu}{\sigma-1} \ln\left\{[1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N]\phi\right\} & \text{(for } i \in \mathcal{N}_A\text{)} \quad (31c) \\ \dfrac{\mu}{\sigma} \dfrac{\delta_N^N[(n+\nu)\phi^2+(n-2)\nu\phi+\nu]}{[1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N]\phi} + \dfrac{\mu}{\sigma-1} \ln\left\{[1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N]\phi\right\} & \text{(for } i \in \mathcal{N}_N \backslash \{l\}\text{).} \quad (31d) \end{cases}$$

By $1 > \delta_A^N > \delta_N^N$, (31b)>(31c)>(31d) always holds, and thus, the condition on whether residing in country i_C is the best for skilled workers of language type $l \in \mathcal{N}_N$ reduces to a comparison between

(31a) and (31b). Then, the condition for skilled workers with language type $l \in \mathcal{N}_N$ is

$$\Delta V^{\mathcal{N}_N}|_{\Lambda=\Lambda_{aggA}} \equiv V_i^l|_{\Lambda=\Lambda_{aggA}, l\in\mathcal{N}_N, i=i_C} - V_i^l|_{\Lambda=\Lambda_{aggA}, l\in\mathcal{N}_N, i=l} > 0. \tag{32}$$

In the second step, we compare the above obtained conditions, (28), (30), and (32), and check which is the strictest condition to satisfy, namely, the final condition for the stability of Λ_{aggA} is

$$\min\{\Delta V^{i_C}|_{\Lambda=\Lambda_{aggA}}, \Delta V^{\mathcal{N}_A\setminus\{i_C\}}|_{\Lambda=\Lambda_{aggA}}, \Delta V^{\mathcal{N}_N}|_{\Lambda=\Lambda_{aggA}}\} > 0. \tag{33}$$

By comparing (28), (30), and (32) under the assumption on δ_i^l, (1), the condition (33) reduces to

$$\Delta V^{\mathcal{N}_N}|_{\Lambda=\Lambda_{aggA}} > 0, \tag{34}$$

which is equivalent to the condition in the text, (16).

Appendix D Comparison of the sustain functions, f_A and f_N

We show that $f_A > f_N$ holds if δ_N^N or δ_N^A is small. In doing so, we derive the condition in which $f_A - f_N > 0$ is guaranteed. By (17) and (20),

$$\begin{aligned}
f_A - f_N &= -\frac{(n+\nu)\phi^2 + [(n-2)\nu - \delta_A^N n(1+\nu)]\phi + \nu}{\sigma[1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N]\phi} + \frac{(n+\nu)\phi^2 + [(n-2)\nu - \delta_N^N n(1+\nu)]\phi + \nu}{\sigma\phi[1+(n_N-1)\delta_N^N + (n-n_N)\delta_N^N]} \\
&> -\frac{(n+\nu)\phi^2 + [(n-2)\nu - \delta_N^N n(1+\nu)]\phi + \nu}{\sigma[1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N]\phi} + \frac{(n+\nu)\phi^2 + [(n-2)\nu - \delta_N^N n(1+\nu)]\phi + \nu}{\sigma\phi[1+(n_N-1)\delta_N^N + (n-n_N)\delta_N^N]} \\
&= \frac{\overbrace{(n+\nu)\phi^2 + [(n-2)\nu - n(1+\nu)\delta_N^N]\phi + \nu}^{\equiv f_1}}{\sigma[1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N]\phi[1+(n_N-1)\delta_N^N + (n-n_N)\delta_N^N]} \\
&\quad \times \left(\underbrace{-(n-n_N)\delta_N^A - (n_N-1)\delta_N^N + (n_A-1)\delta_A^A + (n-n_A)\delta_A^N}_{\equiv f_2}\right),
\end{aligned} \tag{35}$$

where the inequality comes from replacing δ_A^N with δ_N^N in the numerator of the first term[19] and using $\delta_N^N < \delta_A^N$, as well as the denominators are positive. In the following, we first show that $f_2 > 0$.

$$\begin{aligned} f_2 &> -(n-n_N)\delta_N^A - (n_N-1)\delta_N^A + (n_A-1)\delta_A^N + \delta_A^N(n-n_A) \\ &= (n-1)(\delta_A^N - \delta_N^A) \\ &> 0, \end{aligned}$$

where the first inequality comes from replacing δ_N^N with δ_N^A, replacing δ_A^A with δ_A^N, and using $\delta_N^A > \delta_N^N$ and $\delta_A^A > \delta_A^N$, and the second inequality comes from $\delta_A^N > \delta_N^A$. Thus, if $f_1 > 0$, then $f_A - f_N > 0$ holds. Otherwise, $f_A - f_N$ can be negative.

Next we derive the condition under which $f_1 > 0$ holds for two cases: (i) $\delta_N^N \leq \delta^* \equiv \frac{(n-2)\nu}{n(1+\nu)}$ and (ii) $\delta_N^N > \delta^*$.

(i) $\boldsymbol{\delta_N^N \leq \delta^*}$

Because $f_1|_{\phi\to 0} = \nu > 0$, $f_1|_{\phi \to 1} = n(1+\nu)(1-\delta_N^N) > 0$, and the coefficient of ϕ^2, $n+\nu$, is positive, if the axis of symmetry of f_1 (a quadratic function of ϕ) is outside the interval of ϕ, then $f_1 > 0$ in the interval $\phi \in (0,1)$. When $\delta_N^N \leq \delta^*$ the axis of symmetry of f_1, $\phi_1 \equiv \frac{-(n-2)\nu + n(1+\nu)\delta_N^N}{2(n+\nu)} \notin (0,1)$, so that $f_1 > 0$ always holds.

(ii) $\boldsymbol{\delta_N^N > \delta^*}$

In this case, the axis of symmetry $\phi_1 \in (0,1)$, so that f_1 can take its minimum at $\phi = \phi_1 \in (0,1)$. If $f_1|_{\phi=\phi_1} > 0$, then $f_1 > 0$ throughout $\phi \in (0,1)$. Otherwise, f_1 can be negative. Because

$$f_1|_{\phi=\phi_1} > 0 \Leftrightarrow \underbrace{-n(1+\nu)(\delta_N^N)^2 + 2(n-2)\nu(1+\nu)\delta_N^N + \nu[4-(n-4)\nu]}_{\equiv f_1^{\min}} > 0,$$

below we investigate $f_1^{\min}$ and see when it is positive. Specifically, we investigate when the minimum of $f_1^{\min}$ is positive. Because $f_1^{\min}$ is a quadratic function of δ_N^N with a negative coefficient of $(\delta_N^N)^2$, it has a maximum at $\delta_N^N = \frac{(n-2)\nu}{n(1+\nu)}$, which coincides with δ^*. Since the case we consider is $\delta_N^N > \delta^*$, $f_1^{\min}$

[19] There are other possibilities of the replacement of δ_A^N with δ_N^N: (a) the replacement of δ_A^N in the first term and δ_N^N in the second term with δ_N^A; or (b) the replacement of δ_N^N with δ_A^N. Then, by following the same procedure below, we obtain the condition that (a) if $\delta_N^A > \delta^*$ and $\delta_A^N \geq \delta^{**}$, $f_A - f_N$ can be negative, and otherwise, $f_A - f_N > 0$ always holds; or (b) if $\delta_A^N > \delta^*$ and $\delta_A^A \geq \delta^{**}$, $f_A - f_N$ can be negative, and otherwise, $f_A - f_N > 0$ always holds. However, comparing the obtained conditions reported in Proposition 3.1, (a), and (b), the strictest one is that reported in Proposition 3.1, which supports the discussions in the inequality replacement in this manner.

has its maximum at the lower bound, $\delta_N^N \to \delta^*$, and has its minimum at the upper bound, $\delta_N^N \to \delta_N^A$ (remember that $\delta_N^N < \delta_N^A$), and between the interval of $\delta_N^N \in (\delta^*, \delta_N^A)$, $f_1^{\min}$ is monotonically decreasing, so that if $f_1^{\min}|_{\delta_N^N \to \delta_N^A} > 0$, $f_1 > 0$ throughout $\phi \in (0,1)$. Otherwise, $f_1 > 0$ can be negative in the interval $\phi \in (0,1)$. By inspecting the condition $f_1^{\min}|_{\delta_N^N \to \delta_N^A} > 0$ given the restriction $\delta_N^A (> \delta_N^N) > \delta^*$, the condition $f_1^{\min}|_{\delta_N^N \to \delta_N^A} > 0$ is equivalent with the condition $\delta_N^A < \frac{(n-2)\nu + 2\sqrt{(n+\nu)\nu}}{n(1+\nu)} \equiv \delta^{**}$. Then, if $\delta_N^N > \delta^*$ and $\delta_N^A < \delta^{**}$, then $f_1 > 0$ holds.

By summing up the results in (i), if $\delta_N^N \le \delta^*$, then $f_1 > 0$, and in (ii), if $\delta_N^N > \delta^*$ and $\delta_N^A < \delta^{**}$, then $f_1 > 0$, we obtain the following proposition: if $\delta_N^N > \delta^*$ and $\delta_N^A \ge \delta^{**}$, $f_A - f_N$ can be negative, and otherwise, $f_A - f_N > 0$ always holds, which is reported in Proposition 3.1.

Appendix E Impact on f_A given by important parameters

In this appendix, we derive Propositions 3.2-3.5 in order.

Proposition 3.2: f_A and δ_A^N

This is simply obtained from taking the first derivative of f_A with respect to δ_A^N as

$$\frac{\partial f_A}{\partial \delta_A^N} = \frac{(n+\nu)(n-n_A)\phi^2 + [n(n_A-1)(1+\nu)\delta_A^A + n[1+(n-n_A-1)\nu] + 2n_A\nu]\phi + (n-n_A)\nu}{\sigma[1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N]^2\phi} > 0$$

Proposition 3.3: f_A and δ_A^A

By taking the first derivative of f_A with respect to δ_A^A,

$$\frac{\partial f_A}{\partial \delta_A^A} = \frac{n_A - 1}{\sigma[1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N]^2\phi}\left(\underbrace{(n+\nu)\phi^2 + [(n-2)\nu - n(1+\nu)\delta_A^N]\phi + \nu}_{\equiv f_{A1}}\right). \tag{36}$$

Because f_{A1} in (36) is the same as f_1 in (35) appearing in Appendix D except the replacement of δ_N^N in (35) with δ_A^N in (36), and because other terms in (36) are positive, the discussion for obtaining the condition $f_{A1} > 0$ is the same as in Appendix D. Then, by following the discussion done in Appendix D, we obtain the result on the impact of δ_A^A on the sustain function f_A as in Proposition 3.3.

Proposition 3.4: f_A and n_A

By taking the first derivative of f_A with respect to n_A,

$$\frac{\partial f_A}{\partial n_A} = \frac{\delta_A^A - \delta_A^N}{\sigma[1+(n_A-1)\delta_A^A+(n-n_A)\delta_A^N]^2\phi}\left(\underbrace{(n+\nu)\phi^2+[(n-2)\nu-n(1+\nu)\delta_A^N]\phi+\nu}_{\equiv f_{A1}}\right). \tag{37}$$

Because, here again, f_{A1} in (37) is the same as f_1 in (35) appearing in Appendix D except the replacement of δ_N^N in (35) with δ_A^N in (36), and because other terms in (37) are positive by using the assumption that $\delta_A^A > \delta_A^N$, the discussion for obtaining the condition $f_{A1} > 0$ is the same. Then, by following the discussion done in Appendix D, we obtain the result on the impact of n_A on the sustain function f_A as in Proposition 3.4.

Proposition 3.5: f_A and ϕ

To obtain Proposition 3.5, we show that (i) $f_A|_{\phi\to0} < 0$, (ii) $f_A|_{\phi\to1} < 0$, and (iii) f_A has a unique maximum in the interval $\phi \in (0,1)$. Because $f_A|_{\phi\to0} = -\infty$ and $f_A|_{\phi\to1} = -\frac{n(1+\nu)(1-\delta_A^N)}{\sigma[1+(n_A-1)\delta_A^A+(n-n_A)\delta_A^N]} <$ 0, (i) and (ii) are certified. For (iii), we take the first derivative of f_A with respect to ϕ,

$$\frac{\partial f_A}{\partial \phi} = \frac{-(\sigma-1)(n+\nu)\phi^2 - [1+(n_A-1)\delta_A^A+(n-n_A)\delta_A^N]\phi + (\sigma-1)\nu}{\sigma(1-\sigma)[1+(n_A-1)\delta_A^A+(n-n_A)\delta_A^N]\phi}.$$

By solving $\frac{\partial f_A}{\partial \phi} = 0$ for ϕ, it is shown that $\frac{\partial f_A}{\partial \phi} = 0$ has only one root ϕ^* in the interval $\phi \in (0,1)$,

$$\phi^* = \frac{-\sigma[1+(n_A-1)\delta_A^A+(n-n_A)\delta_A^N] + \sqrt{4(\sigma-1)^2(n+\nu)\nu+\sigma^2[1+(n_A-1)\delta_A^A+(n-n_A)\delta_A^N]^2}}{2(\sigma-1)(n+\nu)},$$

such that $\frac{\partial f_A}{\partial \phi} \geq 0$ for $\phi \leq \phi^*$ and $\frac{\partial f_A}{\partial \phi} < 0$ for $\phi > \phi^*$, which implies a unique peak of f_A at $\phi = \phi^*$ in the interval $\phi \in (0,1)$.

From the shape of f_A with respect to ϕ, if $f_A^{\max} \equiv f_A|_{\phi=\phi^*} > 0$, then there exists an interval $(\underline{\phi}, \bar{\phi})$ such that $f_A > 0$ for $\phi \in (\underline{\phi}, \bar{\phi})$ and $f_A \leq 0$ for $\phi \notin (\underline{\phi}, \bar{\phi})$. Also, if $f_A^{\max} \leq 0$, then $f_A \leq 0$ throughout $\phi \in (0,1)$. From this observation, we obtain Proposition 3.5.

Appendix F Social welfare comparisons

(I) Comparison among $\boldsymbol{\Lambda_{aggA}}$, $\boldsymbol{\Lambda_{aggN}}$, and $\boldsymbol{\Lambda_{disp}}$ for skilled workers:

From (22), the average indirect utility of skilled workers for the skilled distribution, Λ_{aggA}, Λ_{aggN}, and

Λ_{disp} are[20]

$$\bar{V}|_{\Lambda=\Lambda_{aggA}} = \mu\left\{\frac{1+\nu}{\sigma} + \frac{\ln[1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N]}{\sigma-1}\right\} \tag{38a}$$

$$\bar{V}|_{\Lambda=\Lambda_{aggN}} = \mu\left\{\frac{1+\nu}{\sigma} + \frac{\ln[1+(n_N-1)\delta_N^N + (n-n_N)\delta_N^A]}{\sigma-1}\right\} \tag{38b}$$

$$\bar{V}|_{\Lambda=\Lambda_{disp}} = \mu\left\{\frac{1+\nu}{\sigma} + \frac{\ln[1+(n-1)\phi]}{\sigma-1}\right\}. \tag{38c}$$

Lemma 4.1: comparison between $\boldsymbol{\Lambda_{aggA}}$ and $\boldsymbol{\Lambda_{disp}}$ for skilled workers

We show $\bar{V}_{\Lambda=\Lambda_{aggA}} > \bar{V}_{\Lambda=\Lambda_{aggN}}$ to prove Lemma 4.1. Taking the difference,

$$\bar{V}_{aggA} - \bar{V}_{aggN} = \frac{\mu}{\sigma-1}\ln\left[\frac{1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N}{1+(n_N-1)\delta_N^N + (n-n_N)\delta_N^A}\right]. \tag{39}$$

Because

$$\begin{aligned}
&1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N - [1+(n_N-1)\delta_N^N + (n-n_N)\delta_N^A] \\
&= (n_A-1)\delta_A^A + (n-n_A)\delta_A^N - [(n_N-1)\delta_N^N + (n-n_N)\delta_N^A] \\
&> (n_A-1)\delta_A^N + (n-n_A)\delta_A^N - [(n_N-1)\delta_N^A + (n-n_N)\delta_N^A] \\
&= (n-1)(\delta_A^N - \delta_N^A) \\
&> 0,
\end{aligned}$$

which implies that the term in $\ln[\cdot]$ in (39) is greater than 1. By this and (39), $\bar{V}_{aggA} - \bar{V}_{aggN} > 0$.

(II) Comparison among $\boldsymbol{\Lambda_{aggA}, \Lambda_{aggN}}$, and $\boldsymbol{\Lambda_{disp}}$ for unskilled workers:

By (24) and that the nominal wage for unskilled workers is one, the average indirect utility of unskilled workers for the skilled distribution, $\Lambda_{aggA}, \Lambda_{aggN}$, and Λ_{disp} are

$$\bar{V}^L|_{\Lambda=\Lambda_{aggA}} = 1 + \frac{\mu}{\sigma-1}\left\{\ln[1+(n_A-1)\delta_A^A + (n-n_A)\delta_A^N] + \frac{n-1}{n}\ln\phi\right\} \tag{40a}$$

$$\bar{V}^L|_{\Lambda=\Lambda_{aggN}} = 1 + \frac{\mu}{\sigma-1}\left\{\ln[1+(n_N-1)\delta_N^N + (n-n_N)\delta_N^A] + \frac{n-1}{n}\ln\phi\right\} \tag{40b}$$

$$\bar{V}^L|_{\Lambda=\Lambda_{disp}} = 1 + \frac{\mu}{\sigma-1}\ln[1+(n-1)\phi]. \tag{40c}$$

[20]Unless necessary in the discussion, the term η is omitted.

Lemma 4.2: comparison between Λ_{aggA}, and Λ_{disp} for unskilled workers

Proving that $\bar{V}^L_{aggA} > \bar{V}^L_{aggN}$ is in the same manner as in the proof of Lemma 4.1, so we omit it.

Proposition 4.2: comparison between Λ_{aggA}, and Λ_{disp} for unskilled workers

We show $\bar{V}^L_{\Lambda=\Lambda_{disp}} > \bar{V}^L_{\Lambda=\Lambda_{aggA}}$ below.

$$\begin{aligned}\bar{V}^L_{disp} - \bar{V}^L_{aggA} &= \frac{\mu}{\sigma-1}\left\{\ln[1+(n-1)\phi] - \frac{n-1}{n}\ln\phi - \ln[1+(n_A-1)\delta^A_A + (n-n_A)\delta^N_A]\right\} \\ &> \frac{\mu}{\sigma-1}\left\{\ln[1+(n-1)\phi] - \frac{n-1}{n}\ln\phi - \ln[1+(n_A-1)\delta^A_A + (n-n_A)\delta^A_A]\right\} \\ &= \frac{\mu}{\sigma-1}\ln\left\{\frac{1+(n-1)\phi}{\phi^{\frac{n-1}{n}}[1+(n-1)\delta^A_A]}\right\}.\end{aligned}$$

Then, it suffices to show that $\frac{1+(n-1)\phi}{\phi^{\frac{n-1}{n}}[1+(n-1)\delta^A_A]} > 1$ to complete the proof. Rewriting $\frac{1+(n-1)\phi}{\phi^{\frac{n-1}{n}}[1+(n-1)\delta^A_A]} > 1$,

$$\delta^A_A < \underbrace{\frac{1}{n-1}\left[\frac{1+(n-1)\phi}{\phi^{\frac{n-1}{n}}} - 1\right]}_{\equiv g^L}. \tag{41}$$

Because g^L is monotonically decreasing with respect to ϕ by taking the first derivative, $\frac{\partial g^L}{\partial\phi} = -\frac{n-1}{n}\phi^{-\frac{2n-1}{n}}(1-\phi) < 0$, g^L takes its minimum at $\phi = 1$, with its minimum value $g^{L\min} \equiv g^L|_{\phi\to 1} = 1$. By $\delta^A_A < 1 = g^{L\min} < g^L$, (41) always holds, which completes the proof.

(III) Comparison among $\Lambda_{aggA}, \Lambda_{aggN}$, and Λ_{disp} for the whole society:

Impact of important parameters related to the Anglosphere:

By taking the first derivatives of g with respect to δ^A_A, δ^N_A, and n_A,

$$\begin{aligned}\frac{\partial g}{\partial\delta^A_A} &= \frac{n(1+\nu)(n_A-1)}{1+(n_A-1)\delta^A_A+(n-n_A)\delta^N_A} > 0 \\ \frac{\partial g}{\partial\delta^N_A} &= \frac{n(1+\nu)(n-n_A)}{1+(n_A-1)\delta^A_A+(n-n_A)\delta^N_A} > 0 \\ \frac{\partial g}{\partial\delta^A_A} &= \frac{n(1+\nu)(\delta^A_A-\delta^N_A)}{1+(n_A-1)\delta^A_A+(n-n_A)\delta^N_A} > 0,\end{aligned}$$

all of which reveal positive impact on the desirability of Λ_{aggA} over Λ_{disp} from the social point of view.

Impact of ϕ :

To obtain Proposition 4.3, we show that (i) $g|_{\phi\to 0} < 0$, (ii) $g|_{\phi\to 1} < 0$, and (iii) g has a unique maximum in the interval $\phi \in (0,1)$. Because $g|_{\phi\to 0} = -\infty$ and $g|_{\phi\to 1} = -n(1+\nu)\ln\left[\frac{n}{1+(n_A-1)\delta^A_A+(n-n_A)\delta^N_A}\right] < 0$,

(i) and (ii) are certified. For (iii), we take the first derivative of g with respect to ϕ,

$$\frac{\partial g}{\partial \phi} = -\frac{(n-1)[(n+\nu)\phi - \nu]}{\phi[1+(n-1)\phi]}.$$

By solving $\frac{\partial g}{\partial \phi} = 0$ for ϕ, it is shown that $\frac{\partial g}{\partial \phi} = 0$ has only one root $\phi = \phi^{W*} \equiv \frac{\nu}{n+\nu}$ in the interval $\phi \in (0,1)$, such that $\frac{\partial g}{\partial \phi} \geq 0$ for $\phi \leq \phi^{W*}$ and $\frac{\partial g}{\partial \phi} < 0$ for $\phi > \phi^{W*}$, which implies a unique peak of g at $\phi = \phi^{W*}$ in the interval $\phi \in (0,1)$.

By the shape of g with respect to ϕ, if $g^{\max} \equiv g|_{\phi=\phi^{W*}} > 0$, then there exists an interval $(\underline{\phi}^W, \bar{\phi}^W)$ such that $g > 0$ for $\phi \in (\underline{\phi}^W, \bar{\phi}^W)$ and $g \leq 0$ for $\phi \notin (\underline{\phi}^W, \bar{\phi}^W)$. Also, if $g^{\max} \leq 0$, then $g \leq 0$ throughout $\phi \in (0,1)$. This observation yields Proposition 4.3.