Is bilingual education desirable in multilingual countries?

Yuki, Kazuhiro

Kyoto University

7 March 2018

Is bilingual education desirable in multilingual countries?

Kazuhiro Yuki*

This version: March 2018
First version: December 2016

Abstract

Many developing countries are populated by multiple ethnic groups who use their own language in daily life and in local business, but have to use a common language in national business and in communications with other groups. In these countries, how much weights should be placed on teaching a local ethnic language and teaching a common language is a critical issue. A similar conflict arises in low-income countries in general between teaching skills that are "practical" and directly useful in local jobs, and teaching academic skills that are important in modern sector jobs.

This paper develops a model to examine these questions theoretically. It is shown that balanced education of the two languages/skills is critical for skill development of those with limited wealth for education. It is also found that the balanced education brings higher earnings net of educational expenditure, *only when* a country has favorable conditions (TFP is reasonably high, and education, in particular, common language education [academic education] is reasonably effective) and *only for* those with adequate wealth. Common-language-only (academic-only) education maximizes net earnings of those with *little* wealth, and, when the country's conditions are not good, maximizes net earnings of *all*. This implies that there exists a *trade-off* between educational and economic outcomes for those with little wealth, and, when the conditions are not good, the trade-off exists for everyone without adequate wealth. Policy implications derived from the results too are discussed.

Keywords: language policy, bilingual education, vocational education, human capital, economic development

JEL classification numbers: I25, I28, J24, O15, O17

*Faculty of Economics, Kyoto University, Yoshida-hommachi, Sakyo-ku, Kyoto, 606-8501, Japan; Phone +81-75-753-3532; E-mail yuki@econ.kyoto-u.ac.jp. Helpful comments from participants at the 2017 Asian Meeting of the Econometric Society are gratefully acknowledged. Financial support from JSPS through Grants-in-Aid for Scientific Research 10197395 is acknowledged.

1 Introduction

Many developing countries, particularly of sub-Saharan Africa, are populated by multiple ethnic groups who use their own language in daily life and in local business, but have to use a common language (typically, the language of the former colonizer) in national business and in communications with other groups. In these countries, how much weights should be placed on teaching a local ethnic language and teaching a common language and which language should be used as a language of instruction of other subjects are critical issues. Acquiring skills to use a local language is less demanding, because the language is a mother tongue and a part of the skills are taught at home, but their uses are limited to a local or ethnic community. By contrast, acquiring skills to use a common language is harder, but the skills are very important in many modern sector jobs.

A similar conflict arises in low-income countries in general, including monolingual countries, between teaching skills that are "practical" and directly useful in local jobs (e.g., farming and related skills in an agrarian community) and teaching academic skills that are not "practical" but are important in jobs involving modern business practice and technology. Acquiring the former skills is less difficult to many students, because they are more familiar and a part of the skills are taught at home, but they are not useful in modern sector jobs.

Students and parents have little choice between local language education and common language education (between "practical" vocational education and academic education) in basic education (primary and lower secondary level), because weights on the two types of education are mostly determined by the government.[1] As for the language policy in sub-Saharan Africa, former French colonies had maintained French-only education and former British colonies had conducted mother tongue education partially, although recently Francophone countries began using local languages in education and many Anglophone countries reduced the weight on local language education (Albaugh, 2007; Heugh, 2011a). Many sub-Saharan African countries, where most students do not proceed to post-basic education, offer vocational subjects, such as agriculture, business, and manual training, in basic education (Atchoarena and Delluc, 2001).

A general consensus among specialists on language and education is that using a local ethnic language as a language of teaching and learning at least in primary education is effective for students to acquire adequate language and non-language skills, and the present language policy in sub-Saharan Africa is overly biased towards common language education, even taking into account the higher publication cost of local language materials (Heugh, 2011b).[2]

By contrast, we know very little what is a desirable combination of the two types of education in terms of future *earnings* and what kind of educational and economic policies should be conducted

[1] The paper is *not* concerned with weights on vocational education and academic education in *post*-basic education, in which students can choose from schools or specialities with different weights.

[2] Works by Vawda and Patrinos (1999) and others find that the publication of local language materials is more costly than that of common language materials, especially when the size of publication is small, in several developing countries. However, Heugh (2011b) argues that local language education at least in primary schooling is more effective than education less reliant on a local language, because other costs depend little on the language of teaching and learning and benefits of local language education on educational outcomes are high.

when *both educational and economic outcomes* of students are taken into account. These questions are important because generally what concerns students and their parents most is future earnings. Despite the recommendation for the increased weight on local language education for skill development by experts, many parents in sub-Saharan Africa are resistant to ethnic language education because they believe that it does not help their children get a job (Albaugh, 2007). Indeed, the economic return to a common language in multilingual countries can be large. For example, Azam, Chin, Prakash (2013) find that the return to speak fluent English, a common language in India, is as large as the return to secondary education and half as large as the return to undergraduate education, after controlling for age, social group, geography, and proxies for ability.

There also seems to be little agreement on desirable weights on vocational contents and academic contents in basic education in poor countries. While the role of academic education for modern sector development cannot be overemphasized, countries such as Ghana increased the weight on vocational contents after a major education reform (Little, 2010). Brock-Utne and Alidou (2011) argue that "practical" education is conductive to the socio-cultural and economic development of students' communities, drawing on episodes from sub-Saharan African countries.

The purpose of this paper is to develop a simple model and examine the above-mentioned important questions theoretically.

Model: In the model, two kinds of "jobs", called *national jobs* and *local jobs*, requiring different types of skills exist and the final good is produced from them. In the real economy, national jobs correspond to many jobs in the modern sector (government and a part of the private sector using modern technology), while local jobs correspond to many jobs in the traditional sector (traditional agriculture, urban informal sector, and household sector).

Each person can expend on education to develop skills required in national jobs and skills required in local jobs. Skills for national jobs correspond to skills to operate a common language and skills for local jobs correspond to skills to operate a local ethnic language in a multilingual country.[3] An alternative interpretation, which would apply to low-income countries in general, is that the former skills are academic skills important in jobs involving modern business practice and technology, while the latter skills are vocational skills directly useful in local jobs.

She can choose the amount of educational spending, but *cannot* choose its allocation over the development of the two types of skills, which is fixed reflecting the fact that weights on the two types of education are mostly determined by the government. The level of skills for local jobs is *positive* without education (i.e. a portion of mother tongue skills and "practical" skills are taught by family members), while the level of skills for national jobs is zero without education.[4]

Although the model without credit constraints too is analyzed, the default setting is that individuals must self-finance education and thus some of them cannot make optimal investment,

[3] Tasks of typical jobs in the modern sector involve communications with people from other parts of the country and thus, in a multiethnic country, with other groups, while tasks of typical jobs in the traditional sector involve communications with mostly locals and thus, in a multiethnic country, with own group.

[4] Further, when the two types of education are interpreted as common language education and local language education, the former is assumed to be less effective in skill development based on empirical findings (footnote 2).

which reflects the fact that, in many developing countries, credit constraints are severe and students must pay for study materials, commuting cost, and others even when public schools do not charge tuitions. After expending on education, each person chooses a job and receives earnings. Because of the assumption that the level of skills for local jobs is positive without education, those with limited wealth choose a local job and those with abundant wealth choose a national job.

Results and policy implications: The paper examines how a change in weights on the two types of education affects educational and job choices and earnings. Main results can be summarized as follows.

First, it is shown that balanced education of skills for national jobs and skills for local jobs is critical for skill development of those with limited wealth: when the allocation of educational spending to the two types of education is very biased, the return to educational investment for local jobs becomes negative, and those who have limited wealth and thus choose a local job do not spend on education.[5] This is consistent with the above-mentioned consensus among experts that using a local ethnic language at least in primary education is effective for skill development.

Second, it is found that balanced education of the two types of skills brings higher earnings net of educational expenditure, *only when* a country has favorable conditions (TFP [total factor productivity] is reasonably high, and education, in particular, the education of skills for national jobs, is reasonably effective) and *only for* those with sufficient wealth. Allocating educational spending completely to the development of skills for *national* jobs (i.e. skills to operate a common language or general academic skills) maximizes net earnings of those with *little* wealth, and, when the country's conditions are not good, maximizes net earnings of *all*. This implies that there always exists a *trade-off* between educational and economic outcomes for those with little wealth, and when the conditions are not good, the trade-off exists for all individuals choosing a local job: under the biased education, their net earnings are highest but their academic performance is lowest. Hence, improved academic performance of students after expansion of the education of skills for local jobs *is not* necessarily a proof that the greater emphasis on the education is desirable.[6]

These results have several policy implications. When the country's conditions are favorable, in order to bring good educational and economic outcomes to all, the government should implement the dual education of an appropriate balance *together with* redistributive policies enabling those with little wealth expend (sufficiently) more on education, such as income transfers and tuition subsidy.[7] Without the latter policies, the very poor lose economically from the implementation of

[5]One might consider the result that the very poor do not spend on education when the allocation of spending is very biased not plausible, since the great majority of students take some education even in poor countries. The difference arises because, for analytical tractability, the model abstracts from motives for attending school other than the investment motive, including consumption motives (joys of studying or attending school) and social motives (pleasure of doing what friends do, pressure from family members or the community to attend school).

[6]It is also found that allocating educational spending completely to the development of skills for national jobs is definitely *better* than introducing education of skills for local jobs *on a small scale*: the latter does not improve academic performance of students from poor families and lowers net earnings of all individuals.

[7]Of course, *given* weights on the two types of education, redistribution toward those without enough wealth for educational investment is desirable in the credit constrained economy, as long as their return to education is positive. Rather, what the statement in the main text asserts is that redistribution toward the very poor is needed

the dual education, because they cannot spend sufficiently enough on education to benefit from it.

By contrast, when the conditions are not good, the government should implement not only the dual education with the redistribution *but also* policies raising the productivity of the economy or the effectiveness of education, in particular, that of skills for national jobs. Only when the latter policies are conducted on a sufficient scale, net earnings become higher under the dual education than under common-language-only (academic-only) education.[8]

Finally, the result that the redistributive policies are essential for the very poor to benefit economically from the dual education gives another justification for governmental support of basic education, in addition to usual rationales based on positive externality, human rights, and among others, in multilingual countries and in low-income countries with large traditional sectors.

Related literature: To the author's knowledge, this paper is the first attempt to examine theoretically how weights on two types of education, common language education and local language education, or academic education and "practical" education in basic education, affect educational investment and net earnings of individuals with different family income. There do exist works examining the issue empirically and works analyzing related issues theoretically.

As mentioned above, researchers in education and linguistics examine the effect of education language policy on academic achievement of students. In economics, there exist a small number of empirical works examining effects on educational and labor market outcomes. Angrist, Chin and Goody (2008) analyze the effect of the policy change in Puerto Rico in 1949, in which Spanish replaced English as the medium of instruction in secondary education, on English skills, and find that the policy change did not lower English skills of the affected students. Laitin, Ramachandran and Walter (2016) examine the effect of an experimental local language schooling program on academic performance of students in Cameroon, and find that effects are sizeable in the short run but fade quickly after students revert to English-medium education. As for labor market outcomes, Angrist and Lavy (1999) find that replacing French with Arabic as the medium of instruction in post-primary education greatly lowered returns to schooling in Morocco. Cappellari and Di Paolo (2015) analyze the effects of the 1983 bilingual education reform in the Catalonia region of Spain, which introduced Catalan alongside Spanish as mediums of instruction, and find positive wage returns to the bilingual education but no effects on employment, hours of work, and occupation.[9]

Pool (1991) develops a model of a multilingual society without intra-group heterogeneity, in which individual earnings do not depend on choice of official language(s), adopting an official language is costly for those whose native language is not official, and when there are multiple official languages, translation among the languages is costly and financed by tax. He shows that there exists efficient and fair choice of official language(s), if appropriate inter-group redistribution is

to *implement* the dual education, i.e. to *choose* intermediate weights.

[8] If the government cannot implement these policies on a sufficient scale for budgetary or other reasons, the dual education with a *smaller* weight on education of skills for local jobs than under when the policies can be conducted might be acceptable: it achieves the higher academic performance of students from relatively poor families than common-language-only (academic-only) education at the relatively small cost of net earnings.

[9] There also exist works examining effects of bilingual education in the U.S., which is monolingual rather than multilingual in business, such as Chin, Daysal, and Imberman (2013) and Lleras-Muney and Shertzer (2015).

conducted. Ginsburgh, Ortuño-Ortín, and Weber (2005) calculate optimal sets of official languages of the EU that minimize the weighted sum of an index of disenfranchisement (the denial of full access to official documents and political process to those whose native languages are not official) and the cost of maintaining official languages. Ortega and Tangerås (2008) develop a model of a society of two language groups without intra-group heterogeneity, in which the politically dominant group determine the type(s) of schools (monolingual school in either language and bilingual school) accessible to each group, individuals decide whether to attend school, and goods are produced from bilateral matching only when pairs speak the same language. They show that the dominant group either choose laissez-faire or restrict access to schools using the language of the dominated, while the dominated prefer schools using their own language.[10]

Main differences between these works and the present work are the following. In the preceding works, language groups are heterogenous, official or education language(s) are chosen from native languages, and individuals within each group are homogenous (except Ginsburgh, Ortuño-Ortín, and Weber (2005), which do not model individual decision-making). By contrast, in the present work, groups are homogenous, common language is not a native language of any group, and individuals are heterogenous with respect to wealth available for education. The present work adopts different settings, because it focuses on developing countries where common language is typically a language of the former colonizer and family wealth is a critical determinant of educational investment, whereas the existing works mainly focus on developed countries.

There also exist theoretical works and case studies examining political aspects of the choice of education languages in multilingual societies. Laitin (1992) describes the choice of common languages in various social domains as a bargaining game among players with different preferences for candidate languages (nationalist leaders, bureaucrats, education specialists, local leaders, parents, among others), and, based on the model and case studies, argues that a stable outcome is multilingual −the colonial language is dominant in administrative, business, educational, and technical domains, and native languages in other domains − in most African countries. Based on case studies of West African countries, Albaugh (2007) argues that the increased use of local languages in education in recent years is largely the consequence of long-held pressures from an alliance of native linguists and NGOs and the changed attitude of the former colonizer, France.

A small number of studies examine empirically the cost effectiveness or economic returns of vocational education relative to academic education in developing countries, although almost all studies focus on *post*-basic education in which students can choose from schools or specialities with different weights on the two types of education (see Eichhorst et al. (2015) and Tan and Nam (2012)

[10] Other works studying issues related to language in economics include the following. Lazear (1999) analyzes a model of adoption of non-native language skills in a multilingual society and empirically examines its implications. Clots-Figueras and Masella (2013) examine effects of the introduction of a bilingual education system in Catalonia on Catalan identity and the propensity to vote for a Catalanist party. Desmet, Ortuño-Ortín, and Wacziarg (2012) empirically examine how measures of linguistic diversity at various levels of linguistic aggregation are associated with civil conflict, growth, public goods provision, and redistribution. Galor, Özak, and Sarid (2018) explore the hypothesis that geographical characteristics of ancestral homelands that were conductive to certain cultural characteristics lead to certain language structures.

for a review). The evidence does not provide a clear picture on the relative effectiveness of the two types of education, which may be partly due to differences in contents of vocational education, such as weights on vocational and academic subjects in vocational schools or tracks, across countries. Based on a theoretical model focusing on tuition-free secondary education of developed countries, Brunello and Giannini (2004) examine the relative efficiency of the stratified system in which students are allocated to academic-only or vocational-only education based on academic ability and the comprehensive system in which all students receive both types of education, and show numerically that neither system unambiguously dominates the other system.

Organization of the paper: Section 2 presents the model and examines educational and job choices of individuals. Section 3 provides preliminary analysis of how a change in weights on the two types of education affects a job choice and earnings of workers, and Section 4, based on results of the previous sections, presents main results of the paper and discusses their policy implications. Section 5 concludes. Appendix A presents some auxiliary results, and Appendix B contains proofs of lemmas and propositions.

2 Model

2.1 Setup

Consider a developing economy in which two kinds of "jobs", called *national jobs* and *local jobs*, requiring different types of skills exist. In the real economy, national jobs correspond to many jobs in the modern sector (government and a part of the private sector using modern technology), whose tasks involve communications with people from other parts of the country and thus, in a multiethnic country, with other groups. Local jobs correspond to many jobs in the traditional sector (traditional agriculture, urban informal sector, and household sector), whose tasks involve communications with mostly locals and thus, in a multiethnic country, with own group.[11,12]

The final good is produced from both types of jobs according to the following technology:

$$Y = A(H_n)^{\alpha}(H_l)^{1-\alpha}, \quad \alpha \in (0,1), \tag{1}$$

where H_n (H_l) is the aggregate human capital of workers in national (local) jobs and A is constant total factor productivity (TFP). The production function implies that the two types of jobs are essential and complementary in the final good production.

Markets are perfectly competitive. From the profit maximization problem of the final good producer, the wage rate per human capital of each type of jobs is given by

$$w_l = (1-\alpha)A\left(\frac{H_n}{H_l}\right)^{\alpha}, \;\; w_n = \alpha A\left(\frac{H_l}{H_n}\right)^{1-\alpha}. \tag{2}$$

[11] Of course, local (national) jobs also include modern (traditional) sector jobs with such task characteristics.

[12] The urban informal sector is a part of the urban economy composed of small-scale businesses supplying basic services (e.g., small shops and vendors selling commodities and meals) and basic manufacturing goods. Even today, sectors or production activities using traditional technology is important in most developing nations (OECD, 2009).

Each person can expend on education to develop skills required in national jobs and skills required in local jobs. Skills required in national jobs correspond to skills to operate a common language (typically, the language of the former colonizer) and skills for local jobs correspond to skills to operate a local ethnic language in a multilingual country.[13] An alternative interpretation, which would apply to low-income countries in general, is that the former skills are academic skills that are not "practical" but are important in jobs involving modern business practice and technology, while the latter skills are "practical" vocational skills directly useful in local jobs, e.g., farming and related skills in an agrarian community.

She can choose the amount of total educational spending e, which largely depends on years of schooling in the real economy, but *cannot* choose its allocation over the development of the two types of skills, which is fixed reflecting the fact that weights on two types of education (common language education and local language education in a multilingual country, and academic contents and vocational contents in a low-income country) are mostly determined by the government in basic education (primary and lower secondary education).[14]

The human capital production functions of the two types of skills are:

$$h_l(e, s_l) = \underline{h}_l + \frac{s_l e}{p}, \tag{3}$$

$$h_n(e, s_l) = \delta \frac{(1 - s_l)e}{p}, \tag{4}$$

$$\text{where } s_l \in [0, 1],\ \underline{h}_l, p, \delta > 0,\ \text{and } e \leq \overline{e}.$$

In the above equations, $\underline{h}_l$ is the level of skills for local jobs when $e = 0$, $s_l \in [0, 1]$ is the proportion of e allocated to the development of skills for local jobs, p is the unit cost of education in terms of the final good (or the reciprocal of productivity of the education of skills for local jobs), and δ is the effectiveness of the education of skills for national jobs relative to that of skills for local jobs. The level of skills for local jobs is positive without education reflecting the fact that they can be developed partly at home (a portion of mother tongue skills and "practical" skills are taught by family members), while the level of skills for national jobs is zero without education. When the two types of education correspond to common language education and local language education, $\delta < 1$ would be reasonable considering the higher cost effectiveness of the latter in skill development (Heugh, 2011b; see footnote 2). The production functions are assumed to be linear to make the model analytically tractable. Because the marginal return to educational investment does not depend on e, i.e. $w_n h_n(1, s_l) - 1 = w_n \delta \frac{1-s_l}{p} - 1$ for national jobs and $w_l \frac{s_l}{p} - 1$ for local jobs, the upper limit $\overline{e}$ is set so that realized e does not become too large for some individuals.[15]

[13] Under this interpretation, the model implicitly assumes that workers of different ethnolinguistic groups with a given type of jobs are perfectly substitutable in production or ethnolinguistic groups are symmetric in every respect.

[14] Many sub-Saharan African countries, where most students do not proceed to post-basic education, offer "practical" vocational subjects, such as agriculture, business, and manual training, in basic education (Atchoarena and Delluc, 2001). Note that the paper is *not* concerned with weights on vocational education and academic education in *post*-basic education, in which students can choose from schools or specialities with different weights.

[15] In particular, the upper limit $\overline{e}$ is needed for demand for e to be bounded when individuals do not face credit

Although the model without credit constraints on educational investment too is analyzed, the default setting is that individuals must self-finance education and thus some of them cannot make optimal investment. This reflects the fact that, in many developing countries, credit constraints on the investment are severe and students must pay for study materials, commuting cost, and supplementary education even when public schools do not charge tuitions.[16] Thus, a person who has wealth (endowment) a available for educational spending can spend at most $e = a$ on education. Let $F(a)$ be the distribution function (and $f(a)$ be the probability density function) of wealth over the population, which is assumed to be continuously differentiable.

After deciding on the amount of education, each person chooses a job and receives earnings, which are spent on the final good for consumption.[17]

2.2 Educational and job choices

Now, educational and job choices and the determination of several endogenous variables are examined in detail.

2.2.1 When education is worthwhile for both types of jobs

First, consider the case in which education is worthwhile, i.e. the return to educational investment is non-negative, for both types of jobs. In this case, those who have wealth (endowment) $a \leq \overline{e}$ spend $e = a$ on education and those with $a > \overline{e}$ spend $e = \overline{e}$ on education. Because both types of jobs are essential in final good production and $h_n(0, s_l) = 0 < h_l(0, s_l) = \underline{h}_l$, there exists $e^+ \in (0, \overline{e}]$ satisfying $w_n h_n(e^+, s_l) = w_l h_l(e^+, s_l)$, and those who spend $e < e^+$ on education choose a local job, whereas, when $e^+ < \overline{e}$, those who spend $e > e^+$ choose a national job, and when $e^+ = \overline{e}$, those who spend $e = \overline{e}$ are indifferent between the two types of jobs. Intuitively speaking, those who have limited wealth and thus cannot spend much on education choose a local job, because a part of skills for such job ($\underline{h}_l$) does not require educational spending. Figure 1 illustrates how educational and job choices and earnings depend on wealth a when $e^+ < \overline{e}$.[18]

The aggregate human capital of workers in local and national jobs, H_l and H_n, are given by

$$\text{When } e^+ < \overline{e},\ H_l(e^+, s_l) = \int_0^{e^+} h_l(e, s_l) f(e) de \text{ and } H_n(e^+, s_l) = \int_{e^+}^{\overline{e}} h_n(e, s_l) f(e) de + [1 - F(\overline{e})] h_n(\overline{e}, s_l), \tag{5}$$

$$\text{When } e^+ = \overline{e},\ H_l(\pi_n, s_l) = \int_0^{\overline{e}} h_l(e, s_l) f(e) de + (1 - \pi_n) [1 - F(\overline{e})] h_l(\overline{e}, s_l) \tag{6}$$

$$\text{and } H_n(\pi_n, s_l) = \pi_n [1 - F(\overline{e})] h_n(\overline{e}, s_l), \tag{7}$$

constraints on educational investment. (The return to educational investment for national jobs *must be* strictly positive, because $w_n h_n(e, s_l) - e = [w_n h_n(1, s_l) - 1] e \geq w_l h_l(0, s_l) = w_l \underline{h}_l$ must hold for those taking a national job.)

[16] Bray and Kwok (2003) briefly review existing studies, which show that the use of private supplementary tutoring is extensive even among primary school students in developing countries.

[17] Wealth (endowment) net of educational spending, $a - e$, too is consumed, which is assumed to yield the same per unit utility as the final good.

[18] When $e^+ = \overline{e}$, $w_l h_l$ intersects with $w_n h_n$ at $a = \overline{e}$, and those with $a \geq \overline{e}$ are indifferent between the two types of jobs.

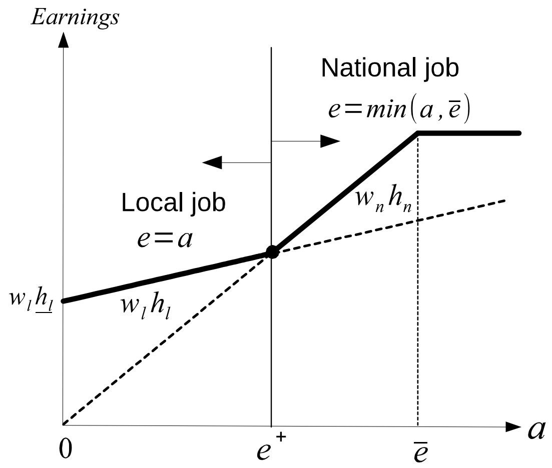


Figure 1: Educational and job choices and earnings when education is worthwhile for all jobs and $e^+ < \overline{e}$

where $\pi_n \in [0, 1]$ is the proportion of individuals choosing a national job among those with wealth $a \geq \overline{e}$ when $e^+ = \overline{e}$ and is equal to (from $w_n h_n(\overline{e}, s_l) = w_l h_l(\overline{e}, s_l)$ and (2))

$$\pi_n = \alpha \left\{ 1 + \frac{\int_0^{\overline{e}} h_l(e, s_l) f(e) de}{[1 - F(\overline{e})] h_l(\overline{e}, s_l)} \right\}. \tag{8}$$

When $e^+ < \overline{e}$, e^+ is determined by

$$w_n h_n(e^+, s_l) = w_l h_l(e^+, s_l) \tag{9}$$

$$\Leftrightarrow \alpha H_l(e^+, s_l) h_n(e^+, s_l) = (1 - \alpha) H_n(e^+, s_l) h_l(e^+, s_l) \quad \text{(from (2))}. \tag{10}$$

2.2.2 When education is not worthwhile for local jobs

Next, consider the case in which education is not worthwhile (the return to educational investment is negative) for local jobs. (Education must always be worthwhile for national jobs, because skills for such jobs cannot be developed without education.) In this case, there exists $e^+ \in (0, \overline{e}]$ satisfying $w_n h_n(e^+, s_l) - e^+ = w_l \underline{h}_l$ and those with wealth $a < e^+$ do not spend on education ($e = 0$) and choose a local job. When $e^+ < \overline{e}$, those with $a > e^+$ choose a national job (those with $a \in (e^+, \overline{e}]$ spend $e = a$ and those with $a > \overline{e}$ spend $e = \overline{e}$), whereas when $e^+ = \overline{e}$, those with $a \geq e^+ = \overline{e}$ are indifferent between the two types of jobs. Figure 2 illustrates how educational and job choices and earnings *net of educational spending* depend on wealth a when $e^+ < \overline{e}$.

In this case, H_l and H_n are given by

$$\text{When } e^+ < \overline{e},\ H_l(e^+, s_l) = F(e^+) \underline{h}_l \text{ and } H_n(e^+, s_l) = \int_{e^+}^{\overline{e}} h_n(e, s_l) f(e) de + [1 - F(\overline{e})] h_n(\overline{e}, s_l), \tag{11}$$

$$\text{When } e^+ = \overline{e},\ H_l(\pi_n, s_l) = \{F(\overline{e}) + (1 - \pi_n)[1 - F(\overline{e})]\} \underline{h}_l \text{ and } H_n(\pi_n, s_l) = \pi_n [1 - F(\overline{e})] h_n(\overline{e}, s_l). \tag{12}$$

When $e^+ < \overline{e}$, e^+ is determined by the following equation with $H_l = H_l(e^+, s_l)$ and $H_n = H_n(e^+, s_l)$

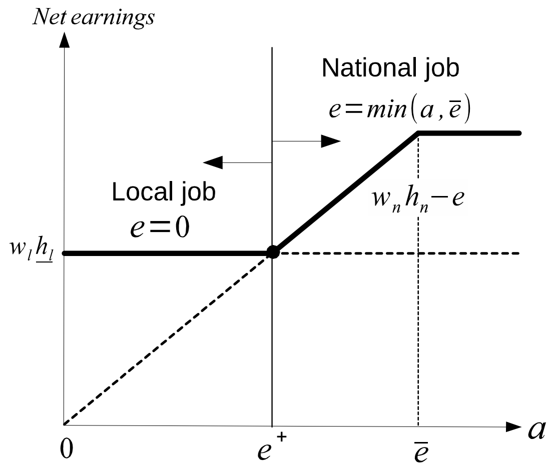


Figure 2: Educational and job choices and earnings net of educational spending when education is not worthwhile for local jobs and $e^+ < \overline{e}$

(equation (11)):

$$w_n h_n(e^+, s_l) - e^+ = w_l h_l(0, s_l) \Leftrightarrow [w_n h_n(1, s_l) - 1] e^+ = w_l \underline{h}_l$$
$$\Leftrightarrow \left[\alpha A \left(\frac{H_l}{H_n}\right)^{1-\alpha} \delta \frac{1-s_l}{p} - 1\right] e^+ = (1-\alpha) A \left(\frac{H_n}{H_l}\right)^{\alpha} \underline{h}_l \quad \text{(from (2))}, \tag{13}$$

while when $e^+ = \overline{e}$, π_n is the solution to the above equation with $H_l = H_l(\pi_n, s_l)$ and $H_n = H_n(\pi_n, s_l)$ (equation (12)).

2.2.3 When is education worthwhile for local jobs and when does $e^+ < \overline{e}$ (or $e^+ = \overline{e}$) hold?

So far, educational and job choices and the determination of several endogenous variables are examined with the sign of the return to education for local jobs and the magnitude relation of e^+ to $\overline{e}$ taken as given. The question is when education is (or is not) worthwhile for local jobs, and when $e^+ < \overline{e}$ (or $e^+ = \overline{e}$) holds. The next proposition provides the answer.

Proposition 1 *Suppose that TFP, A, is not extremely low. Then,*

(i) (a) There exist two critical values of $s_l \in (0, 1)$ at which the return to educational investment for local jobs equals 0, and for s_l smaller (greater) than the lower (higher) critical value, the return is negative and those with wealth $a < e^+$ do not spend on education, while the return is positive and they spend $e = a$ on education for s_l between the critical values.

(b) The lower [higher] critical value of s_l decreases [increases] with A (TFP) and δ (relative effectiveness of education of skills for national jobs) and increases [decreases] with p (unit cost of education) and, when $e^+ < \overline{e}$, $F(\overline{e})$.

(*ii*) $e^+ < \overline{e}$ *holds if* $F(\overline{e})$ *is large or* s_l *is small, and* $e^+ = \overline{e}$ *holds otherwise.*[19] *When* $F(\overline{e}) \leq 1 - \alpha$, $e^+ = \overline{e}$ *always holds.*

The first part of the proposition shows that, when the proportion of educational spending allocated to the development of skills for local jobs, s_l, is very low or *very high*, the return to educational investment for local jobs becomes negative, and those who have limited wealth ($a < e^+$) and thus choose a local job do not spend on education. When s_l is very low, the marginal return $w_l \frac{s_l}{p} - 1$ is negative, because the marginal effect of e on the human capital for local jobs, $\frac{s_l}{p}$, is very small, which dominates high wage rate w_l due to large human capital ratio $\frac{H_n}{H_l}$.[20] When s_l is very high, the return is negative, because w_l becomes very low due to small $\frac{H_n}{H_l}$, which dominates a large marginal effect on the human capital. When s_l is moderate, the return is positive and those who choose a local job spend as much as they can on education, i.e. $e = a$. Higher TFP (total factor productivity), higher effectiveness of education of skills for national jobs (relative to education of skills for local jobs), and lower unit cost of education widen the range of s_l over which education is worthwhile for those choosing a local job, because they raise the wage rate or the marginal effect of educational expenditure on the human capital. The distribution of wealth too affects the sign of the return when $e^+ < \overline{e}$: as the proportion of those who cannot afford the highest level of education $\overline{e}$ falls, i.e. $F(\overline{e})$ falls, the range of s_l with the positive return expands because of higher w_l (see Figure 3).

The result that those choosing a local job do not spend on education when s_l is very low or very high might appear implausible, since the great majority of students take some education even in poor countries. The difference from the real economy arises because, for tractability, the model abstracts from motives for attending school other than the investment motive, including consumption motives (joys of studying or attending school) and social motives (pleasure of doing what friends do, pressure from family members or the community to attend school). The result, however, sheds light on an important source of poor academic performance of students in many developing countries. According to the result, students from modest backgrounds have weak incentive to study and thus perform badly, either because what they learn is mostly irrelevant to their future jobs in the local or ethnic community (when s_l is very low) or because their future earnings are low due to deficient skills of workers in complementary modern sector jobs (when s_l is very high).

The result shows that balanced education of skills for national jobs and skills for local jobs (moderate s_l) is critical for skill development of those with limited wealth. This is consistent with a general consensus among specialists on language and education mentioned in Introduction that using a local ethnic language at least in primary education is effective for students to acquire adequate skills (Heugh, 2011b).

[19] Further, when the return is negative, $e^+ < \overline{e}$ is more likely to hold when A and δ are high or p is low. When the return is positive, $e^+ < \overline{e}$ is more likely to hold when p is low (high) if $\int_0^{\overline{e}} ef(e)de < (>) (1-\alpha)\overline{e}$.

[20] Small s_l implies small H_l and large H_n, since h_l is low and h_n is high and, as shown later, a high proportion of workers choose a national job.

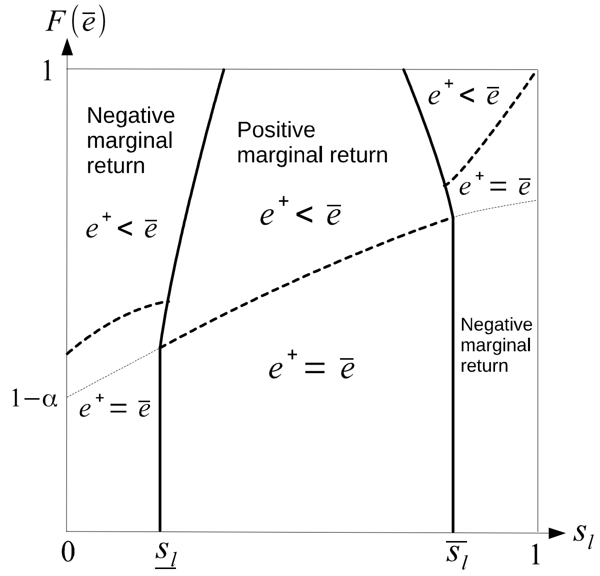


Figure 3: Proposition 1

The second part of the proposition states that $e^+ < \overline{e}$ holds, that is, some individuals who do not have wealth to spend $\overline{e}$ on education get a national job, if the share of such individuals $F(\overline{e})$ is high or the fraction of resources spent on the development of skills for local jobs s_l is low, otherwise, $e^+ = \overline{e}$ holds, that is, everyone who cannot afford $\overline{e}$ takes a local job (see Figure 3).

Figure 3 illustrates the proposition on the $(s_l, F(\overline{e}))$ plane assuming $\int_0^{\overline{e}} ef(e)de < (1-\alpha)\overline{e}$.[21] When s_l is very low or very high, the return to educational investment for local jobs is negative, while the return is positive when s_l is in the intermediate region. (The lower [higher] critical s_l when $e^+ = \overline{e}$, which does not depend on $F(\overline{e})$, is denoted by $\underline{s_l}$ [$\overline{s_l}$] in the figure.) When $e^+ < \overline{e}$, the intermediate region narrows as $F(\overline{e})$ increases. The dividing lines between $e^+ < \overline{e}$ and $e^+ = \overline{e}$ (bold dashed lines) are upward-sloping, thus, for given s_l, $e^+ < \overline{e}$ ($e^+ = \overline{e}$) holds when $F(\overline{e})$ is relatively high (low).[22] Note that $e^+ = \overline{e}$ holds for any s_l when $F(\overline{e}) \leq 1-\alpha$, which is used in later analysis.

3 Preliminary Analysis

This section provides preliminary analysis of how a change in weights on the two types of education affects a job choice and earnings. For ease of presentation, the analysis is conducted mostly *without* taking into account Proposition 1 (Figure 3), which shows that whether education is worthwhile

[21] When $\int_0^{\overline{e}} ef(e)de > (1-\alpha)\overline{e}$, as proved in the proof of (ii) of Proposition 1, $e^+ < \overline{e}$ always holds when the return is positive. When $\int_0^{\overline{e}} ef(e)de = (1-\alpha)\overline{e}$, the dividing line between $e^+ < \overline{e}$ and $e^+ = \overline{e}$ when the return is positive equals $F(\overline{e}) = 1-\alpha$. Main results are unchanged in these cases.

[22] The dividing line when the return is positive is located below the one when the return is negative on the loci for zero return, which is shown in Claim 1 of Appendix A.

for local jobs or not and whether $e^+ < \bar{e}$ or $e^+ = \bar{e}$ holds depend on values of s_l, $F(\bar{e})$, and other parameters. The next section presents main results of the paper by taking into account the proposition, based on the result in this section.

3.1 Effects on a job choice and wage rates

The next lemma examines the effect of s_l on the variables governing a job choice, e^+ when $e^+ < \bar{e}$ and π_n when $e^+ = \bar{e}$.

Lemma 1 *When* $e^+ < \bar{e}$, $\dfrac{de^+}{ds_l} > 0$, *and when* $e^+ = \bar{e}$, $\dfrac{d\pi_n}{ds_l} < 0$.

When the higher proportion of educational spending is allocated to the development of skills for local jobs, the higher fraction of individuals choose a local job, i.e. $\frac{de^+}{ds_l} > 0$ when $e^+ < \bar{e}$ and $\frac{d\pi_n}{ds_l} < 0$ when $e^+ = \bar{e}$. The result can be explained as follows. Remember that a value of the variable governing a job choice, i.e. e^+ when $e^+ < \bar{e}$ and π_n when $e^+ = \bar{e}$, is determined by $w_n h_n(e^+, s_l) = w_l h_l(e^+, s_l)$ when the return to educational investment for local jobs is positive, and by $w_n h_n(e^+, s_l) - e^+ = w_l \underline{h}_l$ when the return is negative. For given e^+ or π_n, an increase in s_l (weakly) raises $h_l(e^+, s_l)$ and lowers $h_n(e^+, s_l)$, which induces some workers to shift from a national job to a local job, while it raises w_n and lowers w_l through a positive effect on the aggregate human capital ratio $\frac{H_l}{H_n}$, which induces the shift in the opposite direction. When $\underline{h}_l > 0$, the former effect dominates and thus the higher fraction of workers choose a local job.

Based on the above lemma, the next lemma examines the effect of s_l on wage rates.

Lemma 2 $\dfrac{dw_n}{ds_l} > 0$ *and* $\dfrac{dw_l}{ds_l} < 0$.

When the higher proportion of spending is allocated to the development of skills for local jobs, the wage rate of national jobs rises and that of local jobs falls. This is because the aggregate human capital ratio $\frac{H_l}{H_n}$ rises due to increased (decreased) human capital of those with a local (national) job and the shift of some workers from a national job to a local job (Lemma 1).

3.2 Effects on earnings

Hence, an increase in s_l raises (lowers) the human capital of workers who choose a local (national) job because of limited (sufficient..) wealth, but lowers (raises) their wage rate. Which effect dominates? The following propositions examine the effect of s_l on earnings based on the lemmas.

The first proposition examines the case in which the return to educational investment for local jobs is negative.

Proposition 2 *When s_l is small or large enough that the return to educational investment for local jobs is negative, earnings decrease with s_l for all individuals.*

When s_l is small or large enough that the return to education for local jobs is negative, as stated in Proposition 1, those who choose a local job due to limited wealth ($a < e^+$) do not spend on education. The proposition shows that, under such situation, earnings of *all individuals* decrease when the greater proportion of educational expenditure is allocated to the development of skills for local jobs. Earnings of those who choose a local job decrease because their wage rate w_l falls due to higher $\frac{H_l}{H_n}$ [because of decreased human capital of those with a national job and the shift of some workers from a national job to a local job] and their human capital remains unchanged at the lowest level, $\underline{h}_l$. Earnings of those choosing a national job fall because higher s_l lowers their human capital, which dominates a positive effect on their wage rate w_n.

The second case to examine is the case in which the return is positive and $e^+ = \bar{e}$ holds. The next proposition summarizes the effect of s_l on earnings for this case, assuming, for ease of presentation, that this case exists for any s_l.

Proposition 3 *Suppose that the return to educational investment for local jobs is positive and* $e^+ = \bar{e}$ *holds.*

(i) $\dfrac{d(w_n h_n)}{ds_l} \gtreqless 0$ *for* $s_l \lesseqgtr s_l^{\star\star} \equiv (1-\alpha) - \alpha\frac{p\underline{h}_l}{\bar{e}}$.[23]

(ii) (a) If $\bar{e} \leq \frac{1+\alpha}{1-\alpha}p\underline{h}_l$, *when* $e > \alpha(p\underline{h}_l + \bar{e})$, $\dfrac{d(w_l h_l)}{ds_l} \gtreqless 0$ *for* $s_l \lesseqgtr s_l^\star(e)$, *where* $s_l^{\star\prime}(e) > 0$ *and* $s_l^\star(e) < s_l^{\star\star}$ *for* $e < \bar{e}$ $(s_l^\star(\bar{e}) = s_l^{\star\star})$,[24] *while when* $e \leq \alpha(p\underline{h}_l + \bar{e})$, $\dfrac{d(w_l h_l)}{ds_l} < 0$ *for any* $s_l > 0$.

(b) Otherwise, when $e \in \left(\frac{\alpha(p\underline{h}_l+\bar{e})}{1+\frac{1}{4p\underline{h}\bar{e}}[(1-\alpha)\bar{e}-(1+\alpha)p\underline{h}_l]^2}, \alpha(p\underline{h}_l+\bar{e})\right)$, $\dfrac{d(w_l h_l)}{ds_l}$ *is negative for* $s_l < s_l^\circ(e) \in (0, s_l^\star(e))$, *where* $s_l^{\circ\prime}(e) < 0$, *positive for* $s_l \in (s_l^\circ(e), s_l^\star(e))$, *and negative for* $s_l > s_l^\star(e)$. *The results when* $e > \alpha(p\underline{h}_l + \bar{e})$ *and when* $e \leq \frac{\alpha(p\underline{h}_l+\bar{e})}{1+\frac{1}{4p\underline{h}\bar{e}}[(1-\alpha)\bar{e}-(1+\alpha)p\underline{h}_l]^2}$ *are same as (a).*[25]

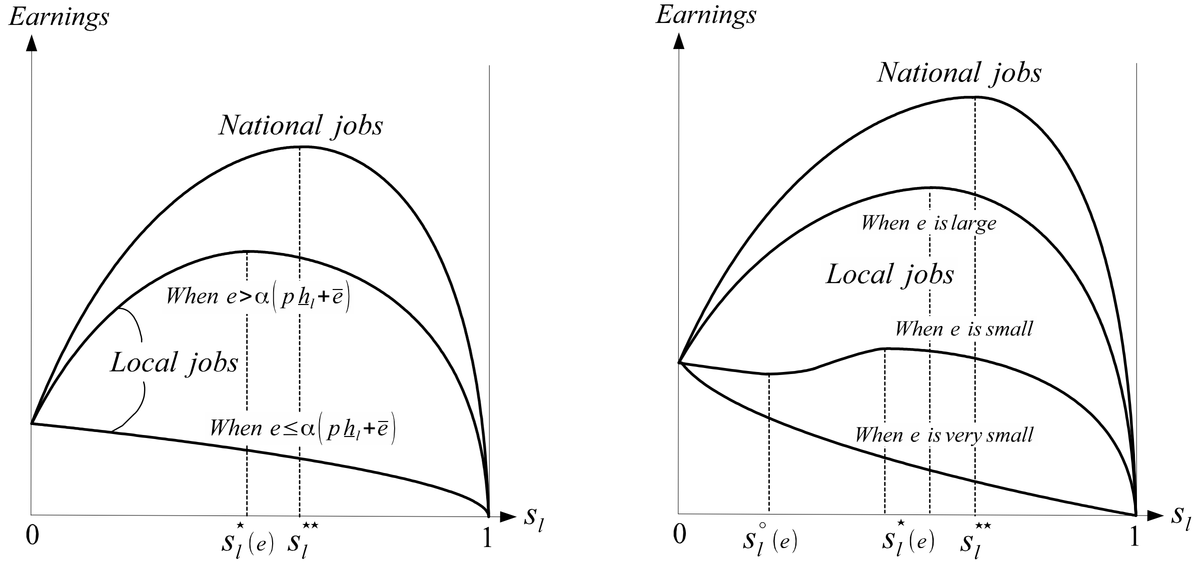
Figure 4 (a) illustrates the proposition when $\bar{e} \leq \frac{1+\alpha}{1-\alpha}p\underline{h}_l$, and Figure 4 (b) illustrates the proposition when $\bar{e} > \frac{1+\alpha}{1-\alpha}p\underline{h}_l$. As mentioned above, the proposition summarizes the effect of s_l on earnings, assuming, for ease of presentation, that the return is positive and $e^+ = \bar{e}$ holds for any s_l, although this is *not* true, as shown in Proposition 1. In particular, when s_l is very low or very large, the return is negative. When main results are presented in the next section, the result of Proposition 1 is taken into account.

In both (a) and (b), earnings of workers with a national job (those with wealth $a \geq e^+ = \bar{e}$) increase with s_l for $s_l < s_l^{\star\star}$, decrease with s_l for $s_l > s_l^{\star\star}$, and are maximized at $s_l = s_l^{\star\star}$. Intuitively speaking, positive expenditure on the education of skills for local jobs maximizes earnings of workers with a national job, because both types of jobs are complementary in production. A more precise explanation is as follows. As mentioned above, an increase in s_l lowers their human capital $h_n(\bar{e}, s_l)$ but raises the wage rate w_n. When s_l is low (high), the latter effect dominates (is dominated by) the former effect, mainly because the aggregate human capital ratio $\frac{H_l}{H_n}$ is relatively low (high) and thus the marginal effect of $\frac{H_l}{H_n}$ on the wage rate is large (small) [see (2)].

[23] $s_l^{\star\star} \equiv (1-\alpha) - \alpha\frac{p\underline{h}_l}{\bar{e}} > 0 \Leftrightarrow \bar{e} > \frac{\alpha}{1-\alpha}p\underline{h}_l$ is assumed thereafter.

[24] $s_l^\star(e)$ ($s_l^\circ(e)$ of (b)) is the greater (smaller) solution of $-e\bar{e}(s_l)^2 + [(1-\alpha)\bar{e} - (1+\alpha)p\underline{h}_l]es_l + [-\alpha(p\underline{h}_l + \bar{e}) + e]p\underline{h}_l = 0$.

[25] When $e = \alpha(p\underline{h}_l + \bar{e})$, $\frac{d(w_l h_l)}{ds_l} \geq (<) 0$ for positive $s_l \leq (>) s_l^\star(e)$ and $\frac{d(w_l h_l)}{ds_l} = 0$ at $s_l = 0$.



(a) When $\bar{e} \leq \frac{1+\alpha}{1-\alpha} p \underline{h}_l$ (b) When $\bar{e} > \frac{1+\alpha}{1-\alpha} p \underline{h}_l$

Figure 4: Effect of s_l on earnings when the return is positive and $e^+ = \bar{e}$ (Proposition 3)

The effect of s_l on earnings of workers with a local job (those with $a \leq e^+ = \bar{e}$) is different depending on the level of $e(= a)$ and in (a) and (b). First, consider case (a) $\bar{e} \leq \frac{1+\alpha}{1-\alpha} p \underline{h}_l$. When the wealth of an individual is high enough that her educational spending satisfies $e > \alpha(p\underline{h}_l + \bar{e})$, her earnings increase (decrease) with s_l for $s_l < (>) s_l^\star(e)$ and are maximized at $s_l = s_l^\star(e)$, where $s_l^\star(e) < s_l^{\star\star}$ and $s_l^{\star\prime}(e) > 0$. The shape of the graph is similar to that of earnings of workers with a national job, although s_l maximizing earnings of workers with a local job is lower.[26] As wealth and thus educational spending e increases, $s_l^\star(e)$ increases and the graph shifts upward. (When $e = e^+ = \bar{e}$, the graph of workers with a local job coincides with that of workers with a national job.) By contrast, when the wealth of an individual is low enough that $e \leq \alpha(p\underline{h}_l + \bar{e})$ holds, her earnings *decrease with* s_l *for any* s_l and thus are maximized at $s_l = 0$.

That is, although higher s_l means the higher proportion of expenditure allocated to the education of skills for local jobs, *no allocation* to the education maximizes earnings of workers with little wealth who choose a local job.

As stated earlier, an increase in s_l raises human capital $h_l(e, s_l)$ but lowers wage rate w_l of workers with a local job. When s_l is low (high), the latter effect tends to be dominated by (dominates) the former effect, mainly because $\frac{H_n}{H_l}$ is relatively high (low) and thus its marginal effect on the wage rate is small (large). Further, the former positive effect through human capital *increases with* e, because an individual with greater wealth and thus educational spending benefits more from the increased weight on the education of skills useful for local jobs. Hence, earnings

[26] This is not necessarily true when $e^+ < \bar{e}$, the next case to examine. Numerical simulations show that, when $e^+ < \bar{e}$, s_l maximizing earnings of workers with a local job can be higher depending on parameters.

of a worker with a local job increase (decrease) with s_l for small (large) s_l and optimal s_l, $s_l^\star(e)$, increases with e, when she has relatively large wealth. By contrast, when she has limited wealth to spend on education, the positive effect through human capital is small and dominated by the negative effect through the wage rate even at $s_l = 0$, thus the earnings are highest at $s_l = 0$.

In case (b) $\overline{e} > \frac{1+\alpha}{1-\alpha} p\underline{h}_l$, the effect of s_l on earnings of workers with a local job is similar to the previous case when wealth (and thus educational spending) is large (i.e. $e > \alpha(p\underline{h}_l + \overline{e})$) and when it is very small (i.e. $e \leq \frac{\alpha(p\underline{h}_l + \overline{e})}{1+\frac{1}{4p\underline{h}\overline{e}}[(1-\alpha)\overline{e}-(1+\alpha)p\underline{h}_l]^2}$), but, in the intermediate range of wealth, their earnings now decrease with s_l for $s_l < s_l^\circ(e) \in (0,\, s_l^\star(e))$ ($s_l^{\circ\prime}(e) < 0$), increase with s_l for $s_l \in (s_l^\circ(e), s_l^\star(e))$, and decrease with s_l for higher s_l (see Figure 4 (b)).[27]

Finally, the case in which the return is positive and $e^+ < \overline{e}$ is examined. The next proposition summarizes analytical results of this case, based on Proposition A1 in Appendix A, which presents detailed results. In the proposition, symbol $e^+(s_l)$ is used to signify the dependence of e^+ on s_l ($e^{+\prime}(s_l) > 0$ from Lemma 1).

Proposition 4 *Suppose that the return to educational investment for local jobs is positive and $e^+ < \overline{e}$ holds.*

(i) (a) If the proportion of those with limited wealth is high enough that $e^+(0) \leq \frac{\alpha p \underline{h}_l}{1-\alpha}$ holds, $\frac{d(w_n h_n)}{ds_l} < 0$ for any s_l.

(b) Otherwise, $\frac{d(w_n h_n)}{ds_l} > (<)\ 0$ for small (large) s_l. s_l satisfying $\frac{d(w_n h_n)}{ds_l} = 0$ is smaller than $s_l^{\star\star}$, the critical s_l when $e^+ = \overline{e}$.

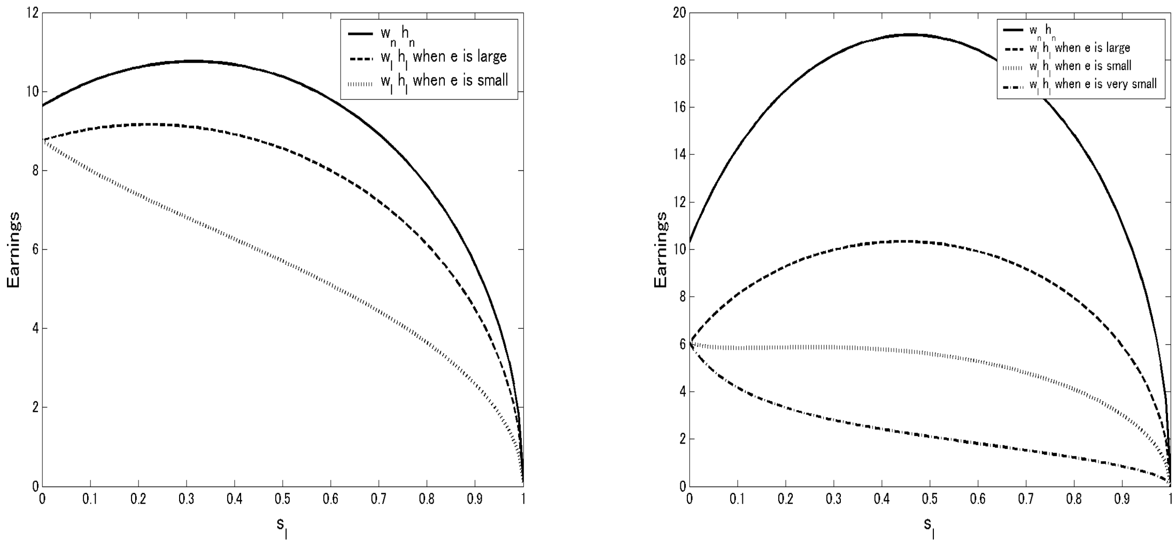
(ii) (a) If $\overline{e} \leq \frac{1+\alpha}{1-\alpha} p\underline{h}_l$, $\frac{d(w_l h_l)}{ds_l} > (<)\ 0$ for small (large) s_l when e is large, while $\frac{d(w_l h_l)}{ds_l} < 0$ for any s_l when e is small. s_l satisfying $\frac{d(w_l h_l)}{ds_l} = 0$ when e is large is greater than $s_l^\star(e)$, the critical s_l when $e^+ = \overline{e}$.

(b) Otherwise, $\frac{d(w_l h_l)}{ds_l}$ is negative for small s_l, positive for middle s_l, and negative for large s_l when e is intermediate, while results when e is small and large are similar to (a). s_l maximizing (minimizing) $w_l h_l$ when e is intermediate is greater (smaller) than $s_l^\star(e)$ ($s_l^\circ(e)$), the critical s_l when $e^+ = \overline{e}$.

(c) The maximum e such that $\frac{d(w_l h_l)}{ds_l} < 0$ holds for any s_l is lower than when $e^+ = \overline{e}$.

Unlike the previous cases, analytical results *cannot* be obtained for some ranges of s_l (see Proposition A1 in Appendix A). The above proposition and numerical simulations, however, suggest that results for workers with a local job are similar to case $e^+ = \overline{e}$: when $\overline{e} \leq \frac{1+\alpha}{1-\alpha} p\underline{h}_l$, $w_l h_l$ increases (decreases) with s_l for small (large) s_l when e is large, and $w_l h_l$ decreases with s_l for any s_l when e is small (like Figure 4 (a) when $e^+ = \overline{e}$); when $\overline{e} > \frac{1+\alpha}{1-\alpha} p\underline{h}_l$, $w_l h_l$ decreases with s_l for small s_l, increases with s_l for middle s_l, and decreases with s_l for large s_l when e is intermediate, while results when e is large and small are similar to the previous case (like Figure 4 (b)). Figure 5

[27] Greater $\overline{e}$ compared to case (a) $\overline{e} \leq \frac{1+\alpha}{1-\alpha} p\underline{h}_l$ implies higher $\frac{H_n}{H_l}$ for given s_l, thus the marginal negative effect of s_l on w_l is smaller. Hence, when $e \leq \alpha(p\underline{h}_l + \overline{e})$ but e is not very small, the earnings increase with s_l for $s_l \in (s_l^-(e), s_l^\star(e))$, unlike the previous case in which they decrease with s_l for any s_l.



(a) When $\overline{e} \leq \frac{1+\alpha}{1-\alpha} p\underline{h}_l$

(b) When $\overline{e} > \frac{1+\alpha}{1-\alpha} p\underline{h}_l$

Figure 5: Effect of s_l on earnings when the return is positive and $e^+ < \bar{e}$ (Proposition 4)

presents a numerical example of the relationship between s_l and earnings of workers with a local job when e is large, small, and very small (in (b) only), and of workers with a national job.[28]

There exist minor differences from case $e^+ = \overline{e}$. First, s_l maximizing earnings of workers with a local job is greater than $s_l^{\star}(e)$, the critical s_l when $e^+ = \overline{e}$.[29,30] (Also, s_l minimizing the earnings when $\overline{e} > \frac{1+\alpha}{1-\alpha} p\underline{h}_l$ and e is intermediate is smaller than $s_l^{\circ}(e)$ when $e^+ = \overline{e}$.) Second, the maximum level of wealth and educational spending such that the earnings decrease with s_l for any s_l (and thus $s_l = 0$ maximizes the earnings) is lower than when $e^+ = \overline{e}$. From Proposition 1 (Figure 3), for given s_l, $e^+ < (=)\overline{e}$ holds when $F(\overline{e})$ is relatively large (small). Hence, these results imply that workers who have limited wealth and thus choose a local job are *more* likely to benefit from the education of skills for local jobs when the proportion of those who cannot afford $\overline{e}$, including themselves, is high, i.e. $e^+ < \overline{e}$, than when the share of such individuals is low, i.e. $e^+ = \overline{e}$.

Results for workers with a national job too are similar to case $e^+ = \overline{e}$, but there exist some

[28] In both (a) and (b), $\alpha = 0.5$, $\underline{h}_l = \delta = 0.5$, the distribution of wealth follows a truncated log normal distribution with maximum 100, mean 6 and variance 10, and the value of e of the earning profile for national jobs equals $\overline{e}$. In (a), $p = 4$, $\overline{e} = 6$, $A = 30$, and values of e of the profiles for local jobs when e is large and small are respectively $1.1\alpha\left(p\underline{h}_l + \overline{e}\right) = 4.4$ and $0.4\Lambda(\overline{e}) = 1.6$, where $\Lambda(\overline{e}) \equiv \frac{\alpha\left(p\underline{h}_l + \overline{e}\right)}{1+\frac{1}{4p\underline{h}\overline{e}}\left[(1-\alpha)\overline{e}-(1+\alpha)p\underline{h}_l\right]^2}$, while in (b), $p = 1$, $\overline{e} = 10$, $A = 10$, and values of e of the profiles when e is large, small, and very small are respectively $\alpha\left(p\underline{h}_l + \overline{e}\right) = 5.25$, $0.87\Lambda(\overline{e}) = 2.4$, and $0.1\Lambda(\overline{e}) = 0.2759$.

[29] Unlike case $e^+ = \overline{e}$, an analytical result cannot be obtained regarding the relationship between s_l maximizing $w_l h_l$ and e. Numerical simulations, however, suggest that the relationship is positive as before.

[30] However, simulations show that, depending on parameters, it is possible that minimum e such that $\frac{d(w_l h_l)}{ds_l} > 0$ for small s_l is greater than e^+ for any s_l (i.e., individuals with such $a = e$ choose a national job), and thus *realized* $w_l h_l$ decreases with s_l *for any* s_l. Simulations also suggest that, in such case, $w_n h_n$ too decreases with s_l for any s_l.

differences. First, if the proportion of those with limited wealth is high enough that $e^+(0) \leq \frac{\alpha p h_l}{1-\alpha}$ holds, $w_n h_n$ decreases with s_l *for any* s_l.[31] Second, when $e^+(0) > \frac{\alpha p h_l}{1-\alpha}$, s_l maximizing $w_n h_n$ is smaller than $s_l^{\star\star}$ when $e^+ = \bar{e}$. Hence, workers who have abundant wealth to choose a national job are *less* likely to benefit from the education of skills for local jobs when $e^+ < \bar{e}$.

4 Main Results

Propositions 2−4 in the previous section examine effects of s_l on earnings for three different cases (the negative return to educational investment for local jobs, the positive return and $e^+ < \bar{e}$, and the positive return and $e^+ = \bar{e}$) separately, by assuming that the economy is in a particular case for any s_l. Proposition 1 (Figure 3) in Section 2.2, however, shows that which of the three cases is realized depends on s_l and other parameters. This section, by taking into account Proposition 1 as well as the results in the previous section, analyzes effects of s_l on earnings *net of educational spending*. Effects on net earnings rather than gross earnings are examined now, because educational spending of an individual could differ depending on which case is realized.

4.1 When $F(\bar{e}) \leq 1 - \alpha$

The next proposition examines effects of s_l on net earnings when the share of individuals who cannot afford the highest level of education $\bar{e}$ is low enough that $F(\bar{e}) \leq 1 - \alpha$ holds, in which case $e^+ = \bar{e}$ always holds (see Figure 3). (In what follows, $\underline{s_l}$ ($\overline{s_l}$) is the lower (higher) critical level of s_l at which the return to educational spending for workers with a local job is 0 when $e^+ = \bar{e}$.)

Proposition 5 *Suppose that $F(\bar{e}) \leq 1 - \alpha$ and thus $e^+ = \bar{e}$ hold.*

(i) If A and δ are small or p is large, net earnings of all workers decrease with s_l.

(ii) Otherwise,

(a) Net earnings of workers with a national job decrease with s_l for $s_l < \underline{s_l}$, increase with s_l for $s_l \in (\underline{s_l}, s_l^{\star\star})$, and decrease with s_l for $s_l > s_l^{\star\star}$. The net earnings are maximized at $s_l = s_l^{\star\star}$ when A and δ are large enough or p is small enough.

(b) Net earnings of those with a local job and wealth above a certain level decrease with s_l for $s_l < \max\{\underline{s_l}, s_l^{\circ}(e)\}$, increase with s_l for $s_l \in (\max\{\underline{s_l}, s_l^{\circ}(e)\}, s_l^{\star}(e))$, and decrease with s_l for $s_l > s_l^{\star}(e)$, while net earnings of those with wealth below the threshold decrease with s_l for any s_l.[32] *Net earnings of the former workers are maximized at $s_l = s_l^{\star}(e)$, when A and δ are large enough or p is small enough.*

(c) The threshold wealth in (b), $\underline{s_l}$, and $\max\{\underline{s_l}, s_l^{\circ}(e)\}$ decrease with A and δ (when $\max\{\underline{s_l}, s_l^{\circ}(e)\} = \underline{s_l}$) and increase with p. $s_l^{\star\star}$ and $s_l^{\star}(e)$ decrease with p.

[31] As mentioned in footnote 26, numerical simulations show that, when $e^+ < \bar{e}$, s_l maximizing earnings of workers with a local job can be higher than s_l maximizing earnings of workers with a national job depending on parameters. Thus, it is possible that $w_n h_n$ decreases with s_l for any s_l and s_l maximizing $w_l h_l$ is positive. Simulations also show that when $e^+(0)$ is low enough, both $w_n h_n$ and *realized* $w_l h_l$ (i.e. $w_l h_l$ with $e \leq e^+$) decrease with s_l for any s_l.

[32] From Proposition 3 (ii), $s_l^{\circ}(e)$ does not exist and thus $\max\{\underline{s_l}, s_l^{\circ}(e)\} = \underline{s_l}$ when $\bar{e} \leq \frac{1+\alpha}{1-\alpha} p \underline{h_l}$.

The first part of the proposition shows that, if A and δ are small or p is large, net earnings of *all workers* decrease with s_l *for any* s_l. This implies that allocating educational spending completely to the education of skills for national jobs ($s_l = 0$) maximizes net earnings of all workers, if TFP is low, the effectiveness of the education of skills for national jobs (relative to the education of skills for local jobs) is low, or the unit cost of education is high.

The result can be explained as follows. When s_l is low or high enough that the return to education for those choosing a local job is negative, i.e. $s_l < \underline{s_l}$ or $s_l > \overline{s_l}$, as shown in Proposition 2, net earnings of all workers decrease with s_l. This is mainly because the greater emphasis on education of skills for local jobs does not affect the human capital of those choosing a local job (unchanged at $\underline{h_l}$) and lowers the human capital of those choosing a national job, which in turn lowers the wage rate of complementary local jobs (see the paragraph just after the proposition for details). When s_l is not at extremes and thus the return is positive, as shown in Proposition 3, net earnings of all workers decrease with s_l when s_l is high, while net earnings of workers with wealth above a threshold increase with s_l when s_l is low. The result when s_l is high (low) is true, roughly because human capital for local (national) jobs is relatively abundant and thus a negative effect of higher s_l on earnings through the decreased human capital for national jobs dominates (is dominated by) a positive effect through the increased human capital for local jobs (see paragraphs after the proposition for details).[33] When TFP is low or education is not effective or costly, i.e. A and δ are small or p is large, the return to education is low for given s_l. Hence, the range of s_l for which the return to education for local jobs is negative expands, i.e. $\underline{s_l}$ increases and $\overline{s_l}$ decreases, and net earnings decrease with s_l for the entire range of s_l for which the return is positive.[34]

By contrast, if A and δ are large or p is small, net earnings of workers with wealth above a certain level, who choose a national job or a local job depending on the level of wealth and s_l, decrease with s_l for small s_l, increase with s_l for intermediate s_l, and decrease with s_l again for large s_l. This result could be understood from Proposition 2 and Figure 4 (Proposition 3). For these workers, an intermediate level of s_l ($s_l^{\star\star}$ for workers with a national job and $s_l^\star(e)$ for workers with a local job and $a = e$) or $s_l = 0$ maximizes net earnings, and, when A and δ are large enough or p is small enough, allocating the expenditure to both types of education is optimal. Finally, net earnings of workers with wealth (and thus educational spending) below the threshold decrease with s_l *for any* s_l. Thus, net earnings of those with little wealth are maximized at $s_l = 0$ *regardless of* the level of TFP and the effectiveness and the cost of education.

As for the latter case, the last part of the proposition shows that, the wealth threshold falls and the range of s_l over which raising s_l increases net earnings of those with wealth above the threshold expands, as A and δ are higher or p is lower. It also shows that, when an intermediate

[33] When s_l is low, net earnings increase with s_l only for workers with wealth above a threshold, because benefits from the increased weight on the education of skills for local jobs are small for those who have limited wealth and thus can spend not much on education.

[34] According to Proposition 3, when the return is positive, net earnings of workers with wealth above the threshold are maximized at $s_l^{\star\star}$ ($s_l^\star(e)$) for those choosing a national (local) job, and their net earnings decrease with s_l for greater s_l. If A and δ are small enough or p is large enough, $\underline{s_l} > s_l^{\star\star}(> s_l^\star(e))$ holds.

level of s_l maximizes their net earnings, the optimal s_l becomes higher as p is lower. These results imply that higher TFP and more effective or less costly education make the greater proportion of people gain from the balanced education of two types of skills, and less costly education makes the greater emphasis on the education of skills for local jobs desirable for these people.

The first part of the proposition implies that the dual education lowers net earnings of workers *irrespective of their wealth*, if TFP is low, the effectiveness of education of skills for national jobs is low, or the unit cost of education is high. This can be seen clearly by considering the case without credit constraints on educational investment, in which all individuals have wealth to finance the highest level of education, i.e. $F(\overline{e}) = 0$.

Corollary 1 *Suppose that everyone has wealth greater than $\overline{e}$, i.e. $F(\overline{e}) = 0$.*

(i) If A and δ are small or p is large, net earnings of all individuals decrease with s_l.

(ii) Otherwise, net earnings of all individuals decrease with s_l for $s_l < \underline{s_l}$, increase with s_l for $s_l \in (\underline{s_l}, s_l^{\star\star})$, and decrease with s_l for $s_l > s_l^{\star\star}$. Their net earnings are maximized at $s_l = s_l^{\star\star}$ when A and δ are large enough or p is small enough.

Without credit constraints, if A and δ are large enough or p is low enough, a balanced allocation of expenditure to both types of education ($s_l = s_l^{\star\star}$) maximizes net earnings of *all workers*, but if not, complete allocation to the education of skills for national jobs remains optimal for all.

4.2 When $F(\overline{e}) > 1 - \alpha$

When the proportion of those who cannot afford $\overline{e}$ is high enough that $F(\overline{e}) > 1 - \alpha$ holds, $e^+ < \overline{e}$ as well as $e^+ = \overline{e}$ could happen depending on levels of s_l and other parameters (see Figure 3). When $e^+ = \overline{e}$, Proposition 5 applies. When $e^+ < \overline{e}$, analytical results on the relation between s_l and net earnings *cannot* be obtained for some ranges of s_l and e (as suggested from Proposition 4 in Section 3 and Proposition A1 in Appendix A). The next proposition, however, shows that main results of this case are similar to the previous case. In what follows, $\underline{s_l}(f)$ ($\overline{s_l}(f)$) denotes the lower (higher) s_l such that the return to educational investment for local jobs is 0 when $e^+ < \overline{e}$. ("f" is to indicate the dependence of their values on the distribution of wealth.)

Proposition 6 *Suppose that $e^+ < \overline{e}$ holds.*[35]

(i) If A and δ are small or p is large, net earnings of all individuals decrease with s_l except at $s_l = \underline{s_l}(f)$, $s_l = \overline{s_l}(f)$, or both, where they increase discontinuously.[36] *The net earnings are maximized at $s_l = 0$ if A and δ are small enough or p is large enough.*

[35] Unlike when $e^+ = \overline{e}$ (Proposition 5), the proposition does not cover intermediate ranges of A, δ, p, and $a = e$.

[36] Net earnings of workers who have abundant wealth and thus choose a national job for any s_l, i.e. $a \geq \min\{e^+(1), \overline{e}\}$, decrease with s_l except at $s_l = \underline{s_l}(f)$; net earnings of workers who have limited wealth and thus choose a local job for any s_l, i.e. $a < e^+(0)$, decrease with s_l except at $s_l = \overline{s_l}(f)$; and net earnings of workers who choose a national (local) job when s_l is small (large) decrease with s_l except at $s_l = \underline{s_l}(f)$, $s_l = \overline{s_l}(f)$, or both.

(*ii*) *Those with wealth below a certain level choose a local job for any s_l and their net earnings decrease with s_l except at $s_l = \overline{s_l}(f)$, where they increase discontinuously, and are maximized at $s_l = 0$.*

(*iii*) *If A and δ are large or p is small enough, net earnings of those with wealth greater than a certain level decrease with s_l for small s_l, increase with s_l for middle s_l, and decrease with s_l for large s_l, and are maximized at $s_l \in (\underline{s_l}(f), \overline{s_l}(f))$.*[37]

The first part of the proposition shows that, if A and δ are small or p is large, net earnings of all individuals decrease with s_l, except at $s_l = \underline{s_l}(f)$, $s_l = \overline{s_l}(f)$, or both, where the return is 0 and the net earnings increase discontinuously, which is different from Proposition 5, but, as before, the net earnings are maximized at $s_l = 0$ if A and δ are small enough or p is large enough.[38] The second part shows that net earnings of individuals with wealth below a certain level decrease with s_l, differently from case $e^+ = \overline{e}$ except at $s_l = \overline{s_l}(f)$, where they increase discontinuously, but as before, their net earnings are maximized at $s_l = 0$. Hence, complete allocation of expenditure to the education of skills for national jobs remains optimal for those with little wealth, and when TFP is low, the effectiveness of education of skills for national jobs is low, or the unit cost of education is high, optimal for others too. Finally, the last part shows that, if A and δ are large enough or p is small enough, net earnings of those with wealth above a certain level are maximized at $s_l \in (\underline{s_l}(f), \overline{s_l}(f))$, implying that the dual education is optimal for them.

4.3 Policy and other implications

What are implications of the above results? This section mainly discusses policy and other implications on the choice of languages of education in a multilingual country.

As mentioned in Introduction, a general consensus among specialists on language and education is that using a local ethnic language at least in primary education is effective for students to acquire adequate skills (Heugh, 2011b). Consistent with this consensus, Proposition 1 implies that balanced education of skills to operate a common language and skills to operate a local language (moderate s_l) is critical for skill development of those who choose a local job due to limited wealth ($a < e^+$).

The results in the previous sections, however, show that the advantage of the balanced dual education in skill development *does not* necessarily translate into higher *economic* returns of the education. Earnings net of educational spending of individuals with little wealth decrease with s_l and common-language-only education brings the highest return to them.[39] Further, when a

[37] Like Proposition 4 in Section 3 and Proposition A1 in Appendix A, on which this result is based, analytical results cannot be obtained for some ranges of s_l. Similar to (i), net earnings increase or decrease discontinuously at $s_l = \underline{s_l}(f)$ and $s_l = \overline{s_l}(f)$. See Lemma A4 in the proof of the proposition for details on directions of change.

[38] Net earnings change discontinuously when the return turns from negative to positive or the other way around, because $\frac{H_n}{H_l}$ (the ratio of human capital of workers with a national job to workers with a local job) changes discontinuously with the discontinuous change in educational investment by the poor. When $e^+ = \overline{e}$ (Proposition 5), by contrast, net earnings change continuously, because a discontinuous change in π_n (the proportion of individuals choosing a national job among those with wealth $a \geq e^+ = \overline{e}$) makes $\frac{H_n}{H_l}$ change continuously.

[39] To be precise, when $e^+ < \overline{e}$, their net earnings decrease with s_l except at $s_l = \overline{s_l}(f)$. The net earnings, however, are always maximized at $s_l = 0$.

country has unfavorable conditions, i.e. the level of TFP is low, the effectiveness of common language education (relative to ethnic language education) in skill development is low, or the unit cost of education is high, net earnings of all individuals decrease with s_l and such education is best for all in terms of the economic outcome.[40] This implies that there exists a *trade-off* between educational and economic outcomes for those with little wealth, and when the conditions are not good, the trade-off exists for everyone choosing a local job: under common-language-only education, their net earnings are highest but their academic performance is poorest.

The results also imply that improved academic performance of students after the expansion of local language education *is not* necessarily a proof that the greater emphasis on the education is desirable. When the initial situation is such that the return to educational investment for local jobs is negative because of very low s_l, the government can change the return to positive by raising s_l appropriately, and can boost educational expenditure (from 0 to a) of those who have limited wealth and end up in a local job (Figure 3). The policy change succeeds in raising their skill. However, it always *lowers* net earnings of the very poor (their *gross* earnings could increase), and when the country's conditions are not good, net earnings of others too.

Further, Proposition 2 implies that introducing local language education *on a small scale* ($s_l < \underline{s_l}$ or $\underline{s_l}(f)$) is *definitely worse* than common-language-only education: the introduction does not improve academic performance of students from poor families and lowers net earnings of all.

What kind of policies should the government implement in order to bring good educational and economic outcomes to everyone? The answer depends on conditions of a country. When it has favorable conditions (TFP is reasonably high, and education, in particular, common language education, is reasonably effective), the government should implement the dual education of an appropriate balance *together with* redistributive policies enabling those with little wealth expend (sufficiently) more on education, such as income transfers, tuition subsidy, and education loans.[41] Without the latter policies, the very poor lose economically from the implementation of the dual education, because they cannot spend sufficiently enough on education to benefit from it.

When the conditions are not good, net earnings of all individuals are highest but academic performance of students from families with limited wealth are lowest under common-language-only education.[42] The dual education of an appropriate weight on local language contents, by contrast, attains the higher academic performance of these students, but at the cost of net earnings of all. The dual education lowers the human capital of workers in national jobs and this has a negative effect on earnings of workers in complementary local jobs as well as on earnings of workers

[40] To be precise, when $e^+ < \bar{e}$, net earnings decrease with s_l except at $s_l = \underline{s_l}(f)$, $s_l = \overline{s_l}(f)$, or both.

[41] Of course, *given* weights on the two types of education (*given* s_l), redistributive policies toward those without abundant wealth ($a < \bar{e}$) are desirable (if the return to education is positive for all jobs or if the policies lift post-transfer wealth of all individuals above $\bar{e}$), because the policies raise their educational spending and net earnings in the credit constrained economy. Rather, what the statement in the main text asserts is that redistribution of a sufficient scale toward the very poor is needed to *implement* the dual education (to *choose* intermediate s_l).

[42] Proposition 1 shows that they do not spend on education because the return to education is negative for local jobs. As mentioned earlier, in the real economy, a great majority of them take some education, most likely because of motives other than the investment motive, which the model abstracts from, but the result shows that their investment motive is very weak, which is consistent with their poor academic performance in the real economy.

in national jobs. The redistributive policies cannot change the lower economic return of the dual education (Corollary 1). On top of the dual education with the redistribution, what the government should conduct is policies raising TFP and the effectiveness of education, in particular, of common language education. If these policies are conducted on a sufficient scale, the positive effect of the dual education on earnings through the increased human capital of workers in local jobs outweighs the negative effect through the decreased human capital of workers in national jobs, and net earnings of everyone become higher under the education. The government, however, may not be able to implement these policies on a sufficient scale for budgetary or other reasons. If this is the case, the dual education with a *smaller* weight on ethnic language contents than under the ideal case, e.g. s_l slightly greater than $\underline{s_l}$ when $e^+ = \bar{e}$, might be acceptable: it achieves the better academic performance of students from poor families than common-language-only education at the relatively small cost of net earnings.

To summarize, although the balanced dual education is essential in raising the educational outcome of students from modest backgrounds, in order to raise their net earnings as well as those of others, other policies must be implemented together, and what kind of complementary policies should be conducted depend on the productivity level of the economy and the cost effectiveness of education, in particular, of common language education.

These implications apply to the issue of "practical" vocational contents versus "non-practical" academic contents in basic education of low-income countries as well, if local (common) language education in the above discussions is replaced with vocational (academic) education. The dual education of academic and vocational contents of an appropriate balance brings the higher academic achievement of students from humble backgrounds than academic-only education, but net earnings of all individuals become higher under such education only if complementary policies are implemented together. If the country's conditions are good, the redistributive policies should be conducted, otherwise, policies raising TFP or the effectiveness of education, in particular, of academic education should be conducted together with the redistribution.

Finally, the result that the redistributive policies are essential for the very poor to benefit economically from the dual education gives another justification for governmental support of basic education, in addition to usual rationales based on positive externality, human rights, and among others, in multilingual countries and in low-income countries with large traditional sectors.

5 Conclusion

Many developing countries are populated by multiple ethnic groups who use their own language in daily life and in local business, but have to use a common language in national business and in communications with other groups. In these countries, how much weights should be placed on teaching a local ethnic language and teaching a common language and which language should be used as a language of instruction of other subjects are critical issues. A similar conflict arises in low-income countries in general between teaching skills that are "practical" and directly useful in local jobs, and teaching academic skills that are important in modern sector jobs.

A general consensus among specialists on language and education is that using a local ethnic language as a language of teaching and learning at least in primary education is effective for students to acquire adequate skills. By contrast, we know very little what is a desirable combination of ethnic language education and common language education in terms of earnings and what kind of educational and economic policies should be conducted when both educational and economic outcomes of students are taken into account. There also seems to be little agreement on desirable weights on "practical" vocational contents and academic contents in basic education.

This paper has developed a simple model to examine these questions. It is shown that balanced education of the two languages/skills is critical for skill development of students with limited wealth for educational investment. It is also found that the balanced education brings higher earnings net of educational expenditure, only when a country has favorable conditions, i.e. productivity is reasonably high, and education, in particular, common language (academic) education, is reasonably effective, and only for those with sufficient wealth. Common-language-only (academic-only) education maximizes net earnings of those with little wealth, and, when the country's conditions are not good, maximizes net earnings of all. This implies that there always exists a trade-off between educational and economic outcomes for those with little wealth, and such trade-off also exists for others choosing a local job, when the conditions are not good.

Several policy implications can be derived from these results. When the country's conditions are favorable, in order to bring good educational and economic outcomes to all, the government should implement the dual education of an appropriate balance together with redistributive policies enabling those with limited wealth expend more on education. By contrast, when the conditions are not good, it should implement not only the dual education with the redistribution but also policies raising the productivity level of the economy, and the effectiveness of education, in particular, of common language (academic) education.

References

[1] Albaugh, Ericka A. (2007), "Language choice in education: a politics of persuasion", *Journal of Modern African Studies* 45(1), pp 1−32.

[2] Angrist, J and Chin, A and Godoy, R. (2008), "Is Spanish-only schooling responsible for the Puerto Rican language gap?" *Journal of Development Economics* 85(1−2), 105−128.

[3] Angrist, J. and Lavy, V. (1997), "The Effect of a change in language of instruction on the returns to schooling in Morocco," *Journal of Labor Economics* 15, S48−S76.

[4] Atchoarena, D. and Delluc, A. (2001), *Revisiting Technical and Vocational Education in sub-Saharan Africa: an Update on Trends, Innovations and Challenges*, UNESCO, IIEP, Paris.

[5] Azam, Mehtabul, Aimee Chin, and Nishith Prakash (2013), "The returns to English-language skills in India," *Economic Development and Cultural Change* 61(2), 335−367.

[6] Bray, M. and P. Kwok (2003), "Demand for private supplementary tutoring: conceptual considerations, and socio-economic patterns in Hong Kong," *Economics of Education Review* 22 (6), 611−620.

[7] Brock-Utne, Birgit and Hassana Alidou (2011), "Active students – learning through a language they master," in A. Ouane and C. Glanz eds., *Optimising Learning, Education and Publishing in Africa: The Language Factor*, the UNESCO Institute for Lifelong Learning and the Association for the Development of Education in Africa.

[8] Brunello, Giorgio, and Massimo Giannini (2004), "Stratified or Comprehensive? The Economic Efficiency of School Design," *Scottish Journal of Political Economy* 51(2), 173−193.

[9] Cappellari, Lorenzo and Antonio Di Paolo (2015), "Bilingual Schooling and Earnings: Evidence from a Language-in-Education Reform," IZA Discussion Paper No. 9431.

[10] Chin, Aimee, N. Meltem Daysal, and Scott A. Imberman (2013) "Impact of bilingual education programs on limited English proficient students and their peers: Regression discontinuity evidence from Texas," *Journal of Public Economics* 107: 63−78.

[11] Clots-Figueras, I. and P. Masella (2013), "Education, Language and Identity," *Economic Journal* 123(570), F332−357.

[12] Desmet, K., I. Ortuño-Ortín, and R. Wacziarg (2012), "The Political Economy of Linguistic Cleavages," *Journal of Development Economics* 97, 322−338.

[13] Eichhorst, Werner, Núria Rodríguez-Planas, Ricarda Schmidl, and Klaus F. Zimmermann (2015), "A Roadmap to Vocational Education and Training Systems Around the World," *ILR Review* 68(2), 314–337.

[14] Galor, Oded and Ömer Özak, and Assaf Sarid (2018), "Geographical Origins of Language Structures," available at SSRN: https://ssrn.com/abstract=3097220.

[15] Ginsburgh, V. A., I. Ortuño-Ortín, and S. Weber (2005), "Disenfranchisement in linguistically diverse societies. The case of the European Union, " *Journal of the European Economic Association* 3(4), 946−965.

[16] Heugh, Kathleen (2011a), "Theory and practice − Language education models in Africa: research, design, decisionmaking, and outcomes", in A. Ouane and C. Glanz eds., *Optimising Learning, Education and Publishing in Africa: The Language Factor*, the UNESCO Institute for Lifelong Learning and the Association for the Development of Education in Africa.

[17] Heugh, Kathleen (2011b), "Cost implications of the provision of mother-tongue and strong bilingual models of education in Africa," in A. Ouane and C. Glanz eds., *Optimising Learning, Education and Publishing in Africa: The Language Factor*, the UNESCO Institute for Lifelong Learning and the Association for the Development of Education in Africa.

[18] Laitin, David D. (1992), *Language Repertoires and State Construction in Africa,* Cambridge University Press.

[19] Laitin, David D., Rajesh Ramachandran, and Stephen L. Walter (2016), "Language of Instruction and Student Learning: Evidence from an Experimental Program in Cameroon," mimeo.

[20] Lazear, E. P. (1999), “Culture and language,” *Journal of Political Economy* 107(6), S95−S126.

[21] Little, Angela W. (2010), "Access to Basic Education in Ghana: politics, policies and progress," CREATE PATHWAYS TO ACCESS Research Monograph No. 42, The Institute of Education, University of London.

[22] Lleras-Muney, Adriana and Allison Shertzer (2015), "Did the Americanization Movement Succeed? An Evaluation of the Effect of English-Only and Compulsory Schooling Laws on Immigrants." *American Economic Journal: Economic Policy* 7(3), 258−90.

[23] OECD (2009), *Is Informal Normal? Towards More and Better Jobs*, OECD, Paris.

[24] Ortega, J. and T. P. Tangerås (2008), “Unilingual versus bilingual education: A political economy analysis,” *Journal of the European Economic Association* 6(5), 1078−1108.

[25] Pool, Jonathan (1991), "The Official Language Problem," *American Political Science Review* 85(2), 495−514.

[26] Tan, Jee-Peng and Yoo-Jeung Joy Nam (2012), "Pre-employment technical and vocational education and training: Fostering relevance, effectiveness, and efficiency," in Almeida, Rita, Jere Behrman, and David Robalino eds., *The right skills for the job?: Rethinking training policies for workers*, World Bank, Washington, DC.

[27] Vawda, Ayesha Yaqub and Harry Anthony Patrinos (1999), "Producing Educational Materials in Local Languages: Costs from Guatemala and Senegal," *International Journal of Educational Development* 19: 287−299.

[28] Wolff, Ekkehard (2011), "Background and history − language politics and planning in Africa", in A. Ouane and C. Glanz eds., *Optimising Learning, Education and Publishing in Africa: The Language Factor*, the UNESCO Institute for Lifelong Learning and the Association for the Development of Education in Africa.

Appendix A: Proposition A1 and Claim 1

This Appendix presents detailed results on effects of s_l on earnings when the return to educational investment for local jobs is positive and $e^+ < \bar{e}$. Proposition 4 in Section 3 provides summarized results. The appendix also presents Claim 1 that is used for drawing Figure 3 in Section 2. Proofs of the proposition and the claim are contained in Appendix C posted on the author's webpage.[43]

Proposition A1 *Suppose that the return to educational investment for local jobs is positive and $e^+ < \bar{e}$ holds.*

[43] The address is http://www.econ.kyoto-u.ac.jp/~yuki/english.html.

(i) (a) If $e^+(0) \le \frac{\alpha p\underline{h}_l}{1-\alpha}$, *which is the case in which the proportion of individuals with limited wealth is high,* $\frac{d(w_n h_n)}{ds_l} < 0$ *for any* s_l.

(b) Otherwise, $\frac{d(w_n h_n)}{ds_l} < 0$ *for* $s_l \ge s_l^\sharp$, *where* $s_l^\sharp \in (0, s_l^{\star\star})$ *satisfies* $s_l^\sharp = (1-\alpha) - \alpha \frac{p\underline{h}_l}{e^+(s_l^\sharp)}$, *and when* $E(e|e<e^+(0)) \equiv \frac{\int_0^{e^+(0)} ef(e)de}{F(e^+(0))} > \max\left\{\frac{\alpha p\underline{h}_l}{1-\alpha}, \frac{\alpha e^+(0)}{1+\alpha}\right\}$, $\frac{d(w_n h_n)}{ds_l} > 0$ *for* $s_l \le s_l^\flat$, *where* $s_l^\flat \in (0, s_l^\sharp)$ *satisfies* $s_l^\flat = (1-\alpha) - \alpha \frac{p\underline{h}_l}{E(e|e<e^+(s_l^\flat))}$.[44]

(ii) (a) If $\overline{e} \le \frac{1+\alpha}{1-\alpha} p\underline{h}_l$, $\frac{d(w_l h_l)}{ds_l} > 0$ *for* $s_l \le s_{l,h}^\triangle(e) \in (s_l^\star(e), s_{l,h}^\triangledown(e))$ *when* $e > \max\{\alpha(p\underline{h}_l + e^+(0)), \Lambda(\overline{e})\}$, $\frac{d(w_l h_l)}{ds_l} < 0$ *for* $s_l \ge s_{l,h}^\triangledown(e)$ *when* $e > \max\{\alpha[p\underline{h}_l + E(e|e<e^+(0))], \Omega(\overline{e})\}$, *and* $\frac{d(w_l h_l)}{ds_l} < 0$ *for any* s_l *when* $e \le \alpha[p\underline{h}_l + E(e|e<e^+(0))]$, *where* $s_{l,h}^\triangle(e)$ *is the greater solution of* $L(s_l) \equiv -ee^+(s_l)s_l^2 + [(1-\alpha)e^+(s_l) - (1+\alpha)p\underline{h}_l]es_l + [-\alpha(p\underline{h}_l + e^+(s_l)) + e]p\underline{h}_l = 0$ *and* $\Lambda(\overline{e}) \equiv \frac{\alpha(p\underline{h}_l + \overline{e})}{1 + \frac{1}{4p\underline{h}_l\overline{e}}[(1-\alpha)\overline{e} - (1+\alpha)p\underline{h}_l]^2}$, *while* $s_{l,h}^\triangledown(e)$ *is the greater solution of* $M(s_l) = 0$, *where* $M(s_l)$ *equals* $L(s_l)$ *with* $e^+(s_l)$ *replaced with* $E(e|e<e^+(s_l))$, *and* $\Omega(\overline{e})$ *equals* $\Lambda(\overline{e})$ *with* $\overline{e}$ *replaced with* $E(e|e<\overline{e})$.

(b) Otherwise, $\frac{d(w_l h_l)}{ds_l}$ *is positive for* $s_l \in \left[\max\{0, s_{l,l}^\triangle(e)\}, s_{l,h}^\triangle(e)\right]$ *when* $e > \max\{\alpha(p\underline{h}_l + e^+(0)), \Lambda(\overline{e})\}$ *(*$s_{l,l}^\triangle(e) < 0$ *when* $e \ge \alpha(p\underline{h}_l + \overline{e})$*), negative for* $s_l \le \max\{0, s_{l,l}^\triangledown(e)\}$ *(*$s_{l,l}^\triangle(e) \in (s_{l,l}^\triangledown(e), s_l^\circ(e))$*) and* $s_l \ge s_{l,h}^\triangledown(e)$ *when* $e \ge \max\{\alpha[p\underline{h}_l + E(e|e<e^+(0))], \Omega(\overline{e})\}$ *(*$s_{l,l}^\triangledown(e) < 0$ *when* $e \ge \alpha[p\underline{h}_l + E(e|e<\overline{e})]$ *and when* $E(e|e<\overline{e}) \le \frac{1+\alpha}{1-\alpha}p\underline{h}_l$*), and negative for any* s_l *when* $E(e|e<e^+(0)) \le \frac{1+\alpha}{1-\alpha}p\underline{h}_l$ *and* $e \le \alpha[p\underline{h}_l + E(e|e<e^+(0))]$ *or when* $e \le \Omega(e^+(0))$, *where* $s_{l,l}^\triangle(e)$ *(*$s_{l,l}^\triangledown(e)$*) is the lower solution of* $L(s_l) = 0$ *(*$M(s_l) = 0$*).*

Claim 1 *Suppose* $\int_0^{\overline{e}} ef(e)de \in (0, (1-\alpha)\overline{e}]$. *As illustrated in Figure 3, on the* $(s_l, F(\overline{e}))$ *plane, the dividing line between* $e^+ < \overline{e}$ *and* $e^+ = \overline{e}$ *when the return to educational investment for local jobs is positive is located below the dividing line when the return is negative on the loci for zero return.*

Appendix B: Proofs of Lemmas and Propositions (Possibly not for publication)

Proof of Proposition 1. (i)(a) Suppose that education is worthwhile for local jobs. Consider case $e^+ < \overline{e}$ first. From (2), (3), (4), and (10), the marginal return to educational investment for local jobs when $e^+ < \overline{e}$ equals

$$w_l \frac{s_l}{p} - 1 = (1-\alpha)A\left[\frac{\alpha h_n(e^+, s_l)}{(1-\alpha)h_l(e^+, s_l)}\right]^\alpha \frac{s_l}{p} - 1 = (1-\alpha)A\left[\frac{\alpha\delta\frac{(1-s_l)e^+}{p}}{(1-\alpha)\left(\underline{h}_l + \frac{s_l e^+}{p}\right)}\right]^\alpha \frac{s_l}{p} - 1. \quad (14)$$

In the above equation,

[44] They are sufficient but not necessary conditions. $\frac{d(w_n h_n)}{ds_l} < (>) 0$ could hold for s_l smaller than $s_l^\sharp$ (greater than $s_l^\flat$) . Similar statements apply to (ii) as well.

$$\frac{d\left(\frac{\frac{(1-s_l)e^+}{p}}{\underline{h}_l+\frac{s_le^+}{p}}\right)}{ds_l}=\frac{\left(\underline{h}_l+\frac{s_le^+}{p}\right)\left(-\frac{e^+}{p}+\frac{1-s_l}{p}\frac{de^+}{ds_l}\right)-\frac{(1-s_l)e^+}{p}\left(\frac{e^+}{p}+\frac{s_l}{p}\frac{de^+}{ds_l}\right)}{\left(\underline{h}_l+\frac{s_le^+}{p}\right)^2}=\frac{-\left(\underline{h}_l+\frac{e^+}{p}\right)\frac{e^+}{p}+\underline{h}_l\frac{1-s_l}{p}\frac{de^+}{ds_l}}{\left(\underline{h}_l+\frac{s_le^+}{p}\right)^2}, \tag{15}$$

where the numerator equals, from (37) in the proof of Lemma 1 below,

$$-\left(\underline{h}_l+\frac{e^+}{p}\right)\frac{e^+}{p}+\underline{h}_l\frac{1-s_l}{p}\frac{de^+}{ds_l}=-\left(\underline{h}_l+\frac{e^+}{p}\right)\frac{e^+}{p}+\underline{h}_l\frac{1-s_l}{p}\frac{\frac{e^+}{p}\left\{(1-\alpha)\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]-\alpha\int_0^{e^+}ef(e)de\right\}}{(1-\alpha)\frac{\underline{h}_l}{e^+}\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]+h_l(e^+,s_l)e^+f(e^+)}$$

$$=\frac{e^+}{p}\frac{-\left(\underline{h}_l+\frac{e^+}{p}\right)\left\{\frac{(1-\alpha)\underline{h}_l}{e^+}\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]+h_l(e^+,s_l)e^+f(e^+)\right\}+\frac{(1-s_l)\underline{h}_l}{p}\left\{(1-\alpha)\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]-\alpha\int_0^{e^+}ef(e)de\right\}}{(1-\alpha)\frac{\underline{h}_l}{e^+}\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]+h_l(e^+,s_l)e^+f(e^+)}$$

$$=-\frac{e^+}{p}\frac{h_l(e^+,s_l)\frac{(1-\alpha)\underline{h}_l}{e^+}\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]+\left(\underline{h}_l+\frac{e^+}{p}\right)h_l(e^+,s_l)e^+f(e^+)+\frac{(1-s_l)\underline{h}_l}{p}\alpha\int_0^{e^+}ef(e)de}{(1-\alpha)\frac{\underline{h}_l}{e^+}\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]+h_l(e^+,s_l)e^+f(e^+)}<0. \tag{16}$$

The sign of the derivative of $w_l\frac{s_l}{p}-1$ with respect to s_l is same as the sign of the following derivative, which, by using the above equations, can be expressed as

$$\frac{d\left(\frac{\frac{(1-s_l)e^+}{p}}{\underline{h}_l+\frac{s_le^+}{p}}(s_l)^{\frac{1}{\alpha}}\right)}{ds_l}=\frac{d\left(\frac{\frac{(1-s_l)e^+}{p}}{\underline{h}_l+\frac{s_le^+}{p}}\right)}{ds_l}(s_l)^{\frac{1}{\alpha}}+\frac{\frac{(1-s_l)e^+}{p}}{\underline{h}_l+\frac{s_le^+}{p}}\frac{1}{\alpha}\frac{(s_l)^{\frac{1}{\alpha}}}{s_l}$$

$$=\frac{(s_l)^{\frac{1}{\alpha}}}{\underline{h}_l+\frac{s_le^+}{p}}\frac{e^+}{p}\left(-\frac{1}{\underline{h}_l+\frac{s_le^+}{p}}\frac{h_l(e^+,s_l)\left\{\frac{(1-\alpha)\underline{h}_l}{e^+}\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]+\left(\underline{h}_l+\frac{e^+}{p}\right)e^+f(e^+)\right\}+\frac{(1-s_l)\underline{h}_l}{p}\alpha\int_0^{e^+}ef(e)de}{(1-\alpha)\frac{\underline{h}_l}{e^+}\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]+h_l(e^+,s_l)e^+f(e^+)}+\frac{1}{\alpha}\frac{1-s_l}{s_l}\right)$$

$$=-\frac{(s_l)^{\frac{1}{\alpha}}}{[h_l(e^+,s_l)]^2}\frac{e^+}{p}\frac{\left(\begin{array}{c}h_l(e^+,s_l)\left\{\frac{(1-\alpha)\underline{h}_l}{e^+}\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]+\left(\underline{h}_l+\frac{e^+}{p}\right)e^+f(e^+)\right\}+\frac{(1-s_l)\underline{h}_l}{p}\alpha\int_0^{e^+}ef(e)de\\-\frac{1}{\alpha}\frac{1-s_l}{s_l}h_l(e^+,s_l)\left\{\frac{(1-\alpha)\underline{h}_l}{e^+}\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]+h_l(e^+,s_l)e^+f(e^+)\right\}\end{array}\right)}{(1-\alpha)\frac{\underline{h}_l}{e^+}\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]+h_l(e^+,s_l)e^+f(e^+)}$$

$$=\frac{(s_l)^{\frac{1}{\alpha}}\frac{e^+}{p}}{[h_l(e^+,s_l)]^2}\frac{h_l(e^+,s_l)\left\{\frac{1-s_l-\alpha s_l}{\alpha s_l}\frac{(1-\alpha)\underline{h}_l}{e^+}\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]+\left(\frac{1-s_l-\alpha s_l}{\alpha s_l}\underline{h}_l+\frac{1-s_l-\alpha}{\alpha}\frac{e^+}{p}\right)e^+f(e^+)\right\}-\frac{(1-s_l)\underline{h}_l}{p}\alpha\int_0^{e^+}ef(e)de}{(1-\alpha)\frac{\underline{h}_l}{e^+}\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]+h_l(e^+,s_l)e^+f(e^+)}. \tag{17}$$

Since $\alpha\left[\int_0^{e^+}(\underline{h}_l+\frac{s_le}{p})f(e)de\right]e^+=(1-\alpha)\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\,\overline{e}\right]h_l(e^+,s_l)$ from (10), (5), (3), and (4), the numerator of the above equation becomes $(s_l)^{\frac{1}{\alpha}}\frac{e^+}{p}$ times

$$\frac{1-s_l-\alpha s_l}{s_l}\underline{h}_l\int_0^{e^+}(\underline{h}_l+\frac{s_l e}{p})f(e)de+\left(\frac{1-s_l-\alpha s_l}{\alpha s_l}\underline{h}_l+\frac{1-s_l-\alpha}{\alpha}\frac{e^+}{p}\right)h_l(e^+,s_l)e^+f(e^+)-\frac{(1-s_l)\underline{h}_l}{p}\alpha\int_0^{e^+}ef(e)de$$
$$=\frac{1}{s_l}\left\{\underline{h}_l\left[(1-s_l-\alpha s_l)\underline{h}_lF(e^+)+(1-s_l-\alpha)\frac{s_l}{p}\int_0^{e^+}ef(e)de\right]+\frac{1}{\alpha}\left[(1-s_l-\alpha s_l)\underline{h}_l+(1-s_l-\alpha)s_l\frac{e^+}{p}\right]h_l(e^+,s_l)e^+f(e^+)\right\}. \tag{18}$$

The following lemma presents the critical result on (18).

Lemma A1 *There exists an $s_l \in (1-\alpha, \frac{1}{1+\alpha})$ such that* (18) *equals zero and the equation is positive* (*negative*) *for lower* (*higher*) s_l.

Proof of Lemma A1. Clearly, (18) is positive for $s_l \leq 1-\alpha$ and negative for $s_l \geq \frac{1}{1+\alpha}$. (18) is positive for s_l greater than $1-\alpha$ and weakly lower than the unique $s_l \in (1-\alpha, \frac{1}{1+\alpha})$ satisfying $(1-s_l-\alpha s_l)\underline{h}_l+(1-s_l-\alpha)s_l\frac{e^+}{p}=0$ too, because, for such s_l, $(1-s_l-\alpha s_l)\underline{h}_l+(1-s_l-\alpha)s_l\frac{e^+}{p}\geq 0$ and $(1-s_l-\alpha s_l)\underline{h}_lF(e^+)+(1-s_l-\alpha)\frac{s_l}{p}\int_0^{e^+}ef(e)de \geq -(1-s_l-\alpha)\frac{s_l}{p}\left[e^+F(e^+)-\int_0^{e^+}ef(e)de\right]>0$, where the former statement is true from

$$\begin{aligned}\frac{d\left[(1-s_l-\alpha s_l)\underline{h}_l+(1-s_l-\alpha)s_l\frac{e^+}{p}\right]}{ds_l} &= -(1+\alpha)\underline{h}_l+(1-2s_l-\alpha)\frac{e^+}{p}+(1-s_l-\alpha)\frac{s_l}{p}\frac{de^+}{ds_l}\\ &< -(1+\alpha)\underline{h}_l+(1-2s_l-\alpha)\frac{e^+}{p}<0 \quad \text{for } s_l>1-\alpha \quad (\text{since } \tfrac{de^+}{ds_l}>0).\end{aligned} \tag{19}$$

Thus, the lemma is proved if the derivative of the expression inside the curly bracket of (18) with respect to s_l is negative for s_l greater than the critical value and lower than $\frac{1}{1+\alpha}$, which equals

$$\begin{aligned}&\underline{h}_l\frac{d\left[(1-s_l-\alpha s_l)\underline{h}_lF(e^+)+(1-s_l-\alpha)\frac{s_l}{p}\int_0^{e^+}ef(e)de\right]}{ds_l}+\frac{1}{\alpha}\frac{d\left[(1-s_l-\alpha s_l)\underline{h}_l+(1-s_l-\alpha)s_l\frac{e^+}{p}\right]}{ds_l}h_l(e^+,s_l)e^+f(e^+)\\ &+\frac{1}{\alpha}\left[(1-s_l-\alpha s_l)\underline{h}_l+(1-s_l-\alpha)s_l\frac{e^+}{p}\right]\frac{d\left[h_l(e^+,s_l)e^+f(e^+)\right]}{ds_l},\end{aligned} \tag{20}$$

where $\frac{d\left[(1-s_l-\alpha s_l)\underline{h}_lF(e^+)+(1-s_l-\alpha)\frac{s_l}{p}\int_0^{e^+}ef(e)de\right]}{ds_l}$

$$\begin{aligned}&=-(1+\alpha)\underline{h}_lF(e^+)+(1-2s_l-\alpha)\frac{1}{p}\int_0^{e^+}ef(e)de+\left[(1-s_l-\alpha s_l)\underline{h}_l+(1-s_l-\alpha)s_l\frac{e^+}{p}\right]f(e^+)\frac{de^+}{ds_l}\\ &<-(1+\alpha)\underline{h}_lF(e^+)+(1-2s_l-\alpha)\frac{1}{p}\int_0^{e^+}ef(e)de<0 \quad \text{for } s_l \text{ greater than the critical value,}\end{aligned} \tag{21}$$

where the first inequality sign is from (19).

Hence, the derivative of (18) is negative if the last term of (20) is negative, which holds unless $f'(e^+)$ is negative and $|f'(e^+)|$ is very large, since $(1-s_l-\alpha s_l)\underline{h}_l+(1-s_l-\alpha)s_l\frac{e^+}{p}<0$ for such s_l. ■

From the lemma, there exists an $s_l \in (1-\alpha, \frac{1}{1+\alpha})$ such that the derivative of the marginal return $w_l\frac{s_l}{p}-1$ with respect to s_l equals zero, and the marginal return increases (decreases) with s_l for s_l smaller (greater) than the critical value. Because the marginal return equals -1 at $s_l=0,1$ from (14), if A is high enough that $w_l\frac{s_l}{p}-1>0$ holds at s_l such that (18) equals zero, there exist two critical values of s_l satisfying $w_l\frac{s_l}{p}-1=0$ and the marginal return is negative for s_l smaller than the lower critical value and greater than the higher one and positive for s_l between them.

When $e^+=\overline{e}$, from (7), (6), and (8),

$$H_l(\overline{e},\pi_n,s_l)=(1-\alpha)\left\{[1-F(\overline{e})]h_l(\overline{e},s_l)+\int_0^{\overline{e}}h_l(e,s_l)f(e)de\right\}, H_n(\overline{e},\pi_n,s_l)=\alpha\left\{[1-F(\overline{e})]h_l(\overline{e},s_l)+\int_0^{\overline{e}}h_l(e,s_l)f(e)de\right\}\frac{h_n(\overline{e},s_l)}{h_l(\overline{e},s_l)}. \tag{22}$$

Thus, from (2)−(4), the marginal return when $e^+ = \overline{e}$ equals

$$w_l\frac{s_l}{p}-1=(1-\alpha)A\left[\frac{\alpha h_n(\overline{e},s_l)}{(1-\alpha)h_l(\overline{e},s_l)}\right]^{\alpha}\frac{s_l}{p}-1=(1-\alpha)A\left[\frac{\alpha\delta\frac{(1-s_l)\overline{e}}{p}(s_l)^{\frac{1}{\alpha}}}{(1-\alpha)\left(\underline{h}_l+\frac{s_l\overline{e}}{p}\right)}\right]^{\alpha}\frac{1}{p}-1. \tag{23}$$

In the above equation,

$$\frac{d\left(\frac{\frac{(1-s_l)\overline{e}}{p}}{\underline{h}_l+\frac{s_l\overline{e}}{p}}(s_l)^{\frac{1}{\alpha}}\right)}{ds_l}=\frac{d\left(\frac{\frac{(1-s_l)\overline{e}}{p}}{\underline{h}_l+\frac{s_l\overline{e}}{p}}\right)}{ds_l}(s_l)^{\frac{1}{\alpha}}+\frac{\frac{(1-s_l)\overline{e}}{p}}{\underline{h}_l+\frac{s_l\overline{e}}{p}}\frac{(s_l)^{\frac{1}{\alpha}}}{\alpha s_l}=\frac{(s_l)^{\frac{1}{\alpha}}\frac{\overline{e}}{p}}{\underline{h}_l+\frac{s_l\overline{e}}{p}}\frac{-\left(\underline{h}_l+\frac{\overline{e}}{p}\right)+\left(\underline{h}_l+\frac{s_l\overline{e}}{p}\right)\frac{1-s_l}{\alpha s_l}}{\underline{h}_l+\frac{s_l\overline{e}}{p}}$$

$$=\frac{(s_l)^{\frac{1}{\alpha}}\frac{\overline{e}}{p}\frac{1}{\alpha s_l}}{\underline{h}_l+\frac{s_l\overline{e}}{p}}\frac{(1-s_l-\alpha s_l)\underline{h}_l+\frac{\overline{e}}{p}s_l(1-s_l-\alpha)}{\underline{h}_l+\frac{s_l\overline{e}}{p}}, \tag{24}$$

which is positive (negative) for s_l smaller (greater) than the critical value satisfying $(1-s_l-\alpha s_l)\underline{h}_l+\frac{\overline{e}}{p}s_l(1-s_l-\alpha)=0$. Hence, the statement is true as in the case of $e^+<\overline{e}$.

(b) The result when $e^+=\overline{e}$ is straightforward from (23).

When $e^+<\overline{e}$, the marginal return depends on e^+ from (14), thus how these exogenous variables affect the return through e^+ must be examined. From (10), (5), (3), and (4), e^+ is a solution to

$$\alpha\int_0^{e^+}\left(\underline{h}_l+\frac{s_le}{p}\right)f(e)dee^+=(1-\alpha)\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\,\overline{e}\right]\left(\underline{h}_l+\frac{s_le^+}{p}\right). \tag{25}$$

Thus, e^+ does not depend on A and δ and the result on these variables is straightforward from (14). e^+ depends negatively on p from (35) in the proof of Lemma 1 and the derivative of the $RHS-LHS$ of the above equation with respect to p, which equals

$$\alpha\int_0^{e^+}\frac{s_le}{p^2}f(e)dee^+-(1-\alpha)\left\{\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\,\overline{e}\right\}\frac{s_le^+}{p^2}=\alpha\frac{s_le^+}{p^2}\left\{\int_0^{e^+}ef(e)de-\frac{\int_0^{e^+}\left(\underline{h}_l+\frac{s_le}{p}\right)f(e)dee^+}{\underline{h}_l+\frac{s_le^+}{p}}\right\}<0. \tag{26}$$

The result on p is clear from (14) and $\frac{de^+}{dp}<0$. The result on $F(\overline{e})$ is from (14) and e^+ being decreasing in $F(\overline{e})$, which is from the fact that the $RHS-LHS$ of (25) decreases with $F(\overline{e})$.

(ii) When the return to educational investment for local jobs is positive, $e^+=\overline{e}$ holds iff (8) satisfies $\pi_n\le 1$, i.e.,

$$\alpha\int_0^{\overline{e}}h_l(e,s_l)f(e)de\le(1-\alpha)[1-F(\overline{e})]h_l(\overline{e},s_l)$$

$$\Leftrightarrow\alpha\int_0^{\overline{e}}\left(\underline{h}_l+\frac{s_l}{p}e\right)f(e)de\le(1-\alpha)[1-F(\overline{e})]\left(\underline{h}_l+\frac{s_l}{p}\overline{e}\right)$$

$$\Leftrightarrow\underline{h}_lF(\overline{e})+\frac{s_l}{p}\left[\alpha\int_0^{\overline{e}}ef(e)de+(1-\alpha)\overline{e}F(\overline{e})\right]\le(1-\alpha)\left(\underline{h}_l+\frac{s_l}{p}\overline{e}\right)$$

$$\Leftrightarrow F(\overline{e})\le\frac{(1-\alpha)\left(\underline{h}_l+\frac{s_l}{p}\overline{e}\right)-\frac{s_l}{p}\alpha\int_0^{\overline{e}}ef(e)de}{\underline{h}_l+\frac{s_l}{p}(1-\alpha)\overline{e}}=1-\alpha\frac{\underline{h}_l+\frac{s_l}{p}\int_0^{\overline{e}}ef(e)de}{\underline{h}_l+\frac{s_l}{p}(1-\alpha)\overline{e}}. \tag{27}$$

When $\int_0^{\overline{e}} ef(e)de > (1-\alpha)\overline{e}$, which implies $F(\overline{e}) > 1-\alpha$, $e^+ = \overline{e}$ cannot hold, because the RHS of (27) decreases with s_l and is smaller than $1-\alpha$. When $\int_0^{\overline{e}} ef(e)de \leq (1-\alpha)\overline{e}$, the RHS weakly increases with s_l, hence $e^+ = \overline{e}$ could hold and $e^+ = \overline{e}$ always when $F(\overline{e}) \leq 1-\alpha$. The rest of the statement is straightforward from (27) (note that $F(\overline{e})$ raises $\int_0^{\overline{e}} ef(e)de$ and thus lowers the RHS).

When the return is negative, from (13), (4), and (12), $w_n h_n(\overline{e}, s_l) - \overline{e} = w_l \underline{h}_l$ can be expressed as

$$\left[\alpha A\left(\frac{\{F(\overline{e})+(1-\pi_n)[1-F(\overline{e})]\}\underline{h}_l}{\pi_n[1-F(\overline{e})]h_n(\overline{e},s_l)}\right)^{1-\alpha}\delta\frac{(1-s_l)}{p}-1\right]\overline{e} = (1-\alpha)A\left(\frac{\pi_n[1-F(\overline{e})]h_n(\overline{e},s_l)}{\{F(\overline{e})+(1-\pi_n)[1-F(\overline{e})]\}\underline{h}_l}\right)^{\alpha}\underline{h}_l. \tag{28}$$

$\pi_n \in (0,1]$ satisfying the above equation exists, that is, $e^+ = \overline{e}$ holds iff the LHS of the above equation is weakly smaller than the RHS at $\pi_n = 1$:

$$\left[\alpha A\left(\frac{F(\overline{e})\underline{h}_l}{[1-F(\overline{e})]h_n(\overline{e},s_l)}\right)^{1-\alpha}\delta\frac{(1-s_l)}{p}-1\right]\overline{e} \leq (1-\alpha)A\left(\frac{[1-F(\overline{e})]h_n(\overline{e},s_l)}{F(\overline{e})\underline{h}_l}\right)^{\alpha}\underline{h}_l$$

$$\Leftrightarrow \left[\alpha A\left(\frac{F(\overline{e})\underline{h}_l}{[1-F(\overline{e})]\overline{e}}\right)^{1-\alpha} - \left(\delta\frac{1-s_l}{p}\right)^{-\alpha}\right]\overline{e} \leq (1-\alpha)A\left(\frac{[1-F(\overline{e})]\overline{e}}{F(\overline{e})\underline{h}_l}\right)^{\alpha}\underline{h}_l$$

$$\Leftrightarrow \left(\delta\frac{1-s_l}{p}\right)^{-\alpha} \geq A\left(\frac{[1-F(\overline{e})]\overline{e}}{F(\overline{e})\underline{h}_l}\right)^{\alpha}\frac{\underline{h}_l}{[1-F(\overline{e})]\overline{e}}[F(\overline{e})-(1-\alpha)]$$

$$\Leftrightarrow \frac{F(\overline{e})-(1-\alpha)}{[F(\overline{e})]^{\alpha}[1-F(\overline{e})]^{1-\alpha}}A\left(\delta\frac{1-s_l}{p}\right)^{\alpha}\left(\frac{\underline{h}_l}{\overline{e}}\right)^{1-\alpha} \leq 1. \tag{29}$$

Clearly, the condition is satisfied when s_l is high and when $F(\overline{e}) \leq 1-\alpha$. It holds when $F(\overline{e})$ is low because the derivative of the first part of the LHS of (29) with respect to $F(\overline{e})$ equals

$$\frac{[F(\overline{e})]^{\alpha}[1-F(\overline{e})]^{1-\alpha}f(\overline{e})\left\{1-[F(\overline{e})-(1-\alpha)]\left(\frac{\alpha}{F(\overline{e})}-\frac{1-\alpha}{1-F(\overline{e})}\right)\right\}}{\left\{[F(\overline{e})]^{\alpha}[1-F(\overline{e})]^{1-\alpha}\right\}^2} = \frac{[F(\overline{e})]^{\alpha}[1-F(\overline{e})]^{1-\alpha}f(\overline{e})\left\{1-\frac{[F(\overline{e})-(1-\alpha)][\alpha-F(\overline{e})]}{F(\overline{e})[1-F(\overline{e})]}\right\}}{\left\{[F(\overline{e})]^{\alpha}[1-F(\overline{e})]^{1-\alpha}\right\}^2}$$

$$= \frac{[F(\overline{e})]^{\alpha}[1-F(\overline{e})]^{1-\alpha}f(\overline{e})\alpha(1-\alpha)}{\left\{[F(\overline{e})]^{\alpha}[1-F(\overline{e})]^{1-\alpha}\right\}^3} > 0. \tag{30}$$

■

Proof of Lemma 1. [When the return to educational investment for local jobs is positive] When $e^+ < \overline{e}$, by totally differentiating (10), one obtains

$$\left[\alpha H_l(e^+,s_l)\frac{\partial h_n(e^+,s_l)}{\partial e^+} - (1-\alpha)H_n(e^+,s_l)\frac{\partial h_l(e^+,s_l)}{\partial e^+} + \alpha\frac{\partial H_l(e^+,s_l)}{\partial e^+}h_n(e^+,s_l) - (1-\alpha)\frac{\partial H_n(e^+,s_l)}{\partial e^+}h_l(e^+,s_l)\right]de^+$$
$$+\left[\alpha H_l(e^+,s_l)\frac{\partial h_n(e^+,s_l)}{\partial s_l} - (1-\alpha)H_n(e^+,s_l)\frac{\partial h_l(e^+,s_l)}{\partial s_l} + \alpha\frac{\partial H_l(e^+,s_l)}{\partial s_l}h_n(e^+,s_l) - (1-\alpha)\frac{\partial H_n(e^+,s_l)}{\partial s_l}h_l(e^+,s_l)\right]ds_l = 0, \tag{31}$$

where

$$\frac{\partial H_l(e^+,s_l)}{\partial e^+} = h_l(e^+,s_l)f(e^+) > 0, \frac{\partial H_n(e^+,s_l)}{\partial e^+} = -h_n(e^+,s_l)f(e^+) < 0. \tag{32}$$

$$\frac{\partial H_l(e^+,s_l)}{\partial s_l} = \int_0^{e^+}\frac{\partial h_l(e,s_l)}{\partial s_l}f(e)de = \frac{1}{p}\int_0^{e^+}ef(e)de > 0. \tag{33}$$

$$\frac{\partial H_n(e^+,s_l)}{\partial s_l} = \int_{e^+}^{\overline{e}}\frac{\partial h_n(e,s_l)}{\partial s_l}f(e)de + [1-F(\overline{e})]\frac{\partial h_n(\overline{e},s_l)}{\partial s_l} = -\frac{\delta}{p}\left\{\int_{e^+}^{\overline{e}}ef(e)de + [1-F(\overline{e})]\overline{e}\right\} < 0. \tag{34}$$

In (31), the term of de^+ equals

$$\alpha H_l(e^+,s_l)\frac{\partial h_n(e^+,s_l)}{\partial e^+} - (1-\alpha)H_n(e^+,s_l)\frac{\partial h_l(e^+,s_l)}{\partial e^+} + h_n(e^+,s_l)h_l(e^+,s_l)f(e^+)$$
$$= \frac{1}{e^+}\left\{\alpha H_l(e^+,s_l)h_n(e^+,s_l) - (1-\alpha)H_n(e^+,s_l)\left[h_l(e^+,s_l) - \underline{h}_l\right]\right\} + h_n(e^+,s_l)h_l(e^+,s_l)f(e^+)$$
$$= \frac{1}{e^+}(1-\alpha)H_n(e^+,s_l)\underline{h}_l + h_n(e^+,s_l)h_l(e^+,s_l)f(e^+) \text{ (from (10))}$$
$$= \delta\frac{1-s_l}{p}\left\{(1-\alpha)\frac{h_l}{e^+}\left[\int_{e^+}^{\overline{e}} ef(e)de + [1-F(\overline{e})]\overline{e}\right] + e^+h_l(e^+,s_l)f(e^+)\right\} > 0, \tag{35}$$

The terms of ds_l equals

$$\alpha H_l(e^+,s_l)\frac{\partial h_n(e^+,s_l)}{\partial s_l} - (1-\alpha)H_n(e^+,s_l)\frac{\partial h_l(e^+,s_l)}{\partial s_l}$$
$$+\alpha h_n(e^+,s_l)\int_0^{e^+}\frac{\partial h_l(e,s_l)}{\partial s_l}f(e)de - (1-\alpha)\left[\int_{e^+}^{\overline{e}}\frac{\partial h_n(e,s_l)}{\partial s_l}f(e)de + [1-F(\overline{e})]\frac{\partial h_n(\overline{e},s_l)}{\partial s_l}\right]h_l(e^+,s_l)$$
$$=\frac{1}{p}\left\{-e^+\left[\alpha H_T(e^+,s_l)\delta+(1-\alpha)H_n(e^+,s_l)\right]+\alpha h_s(e^+,s_l)\int_0^{e^+}ef(e)de+(1-\alpha)\delta\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]h_l(e^+,s_l)\right\}$$
$$=\frac{1}{p}\left\{\begin{array}{c} -\frac{e^+(1-\alpha)H_n(e^+,s_l)}{h_n(e^+,s_l)}\left[\delta h_l(e^+,s_l)+h_n(e^+,s_l)\right] \\ +\alpha h_s(e^+,s_l)\int_0^{e^+}ef(e)de+(1-\alpha)\delta\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]h_l(e^+,s_l)\end{array}\right\} \text{ (from (10))}$$
$$=\frac{1}{p}\left\{\begin{array}{c} -(1-\alpha)\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]\left[\delta h_l(e^+,s_l)+h_n(e^+,s_l)\right] \\ +\alpha h_s(e^+,s_l)\int_0^{e^+}ef(e)de+(1-\alpha)\delta\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]h_l(e^+,s_l)\end{array}\right\} \text{ (from (5) and (4))}$$
$$= -\frac{h_n(e^+,s_l)}{p}\left\{(1-\alpha)\left[\int_{e^+}^{\overline{e}}ef(e)de + [1-F(\overline{e})]\overline{e}\right] - \alpha\int_0^{e^+}ef(e)de\right\} < 0, \tag{36}$$

where the last inequality holds because

$$\alpha\left[\underline{h}_l\int_0^{e^+}f(e)de+\frac{s_l}{p}\int_0^{e^+}ef(e)de\right]e^+ = (1-\alpha)\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]\left(\underline{h}_l + s_l\frac{e^+}{p}\right)$$
$$\Leftrightarrow \left\{\alpha F(e^+)e^+-(1-\alpha)\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]\right\}\underline{h}_l = s_l\frac{e^+}{p}\left\{(1-\alpha)\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]-\alpha\int_0^{e^+}ef(e)de\right\} \text{ from (10)}$$

and thus $sign\left\{(1-\alpha)\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]-\alpha\int_0^{e^+}ef(e)de\right\}=sign\left\{\alpha F(e^+)e^+-(1-\alpha)\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]\right\}$.

Hence,

$$\frac{de^+}{ds_l} = \frac{\frac{h_n(e^+,s_l)}{p}\left\{(1-\alpha)\left[\int_{e^+}^{\overline{e}}ef(e)de + [1-F(\overline{e})]\overline{e}\right] - \alpha\int_0^{e^+}ef(e)de\right\}}{\frac{1}{e^+}(1-\alpha)H_n(e^+,s_l)\underline{h}_l + h_n(e^+,s_l)h_l(e^+,s_l)f(e^+)}$$
$$= \frac{\frac{e^+}{p}\left\{(1-\alpha)\left[\int_{e^+}^{\overline{e}}ef(e)de + [1-F(\overline{e})]\overline{e}\right] - \alpha\int_0^{e^+}ef(e)de\right\}}{(1-\alpha)\frac{h_l}{e^+}\left[\int_{e^+}^{\overline{e}}ef(e)de + [1-F(\overline{e})]\overline{e}\right] + h_l(e^+,s_l)e^+f(e^+)} > 0. \tag{37}$$

When $e^+ = \overline{e}$, from (8),

$$\frac{d\pi_n}{ds_l} = \alpha\frac{h_l(\overline{e},s_l)\int_0^{\overline{e}}\frac{\partial h_l(e,s_l)}{\partial s_l}f(e)de - \frac{\partial h_l(\overline{e},s_l)}{\partial s_l}\int_0^{\overline{e}}h_l(e,s_l)f(e)de}{[1-F(\overline{e})]\left[h_l(\overline{e},s_l)\right]^2} = \alpha\frac{\frac{h_l}{p}\left(\int_0^{\overline{e}}ef(e)de - \overline{e}\int_0^{\overline{e}}f(e)de\right)}{[1-F(\overline{e})]\left[h_l(\overline{e},s_l)\right]^2} < 0. \tag{38}$$

[When the return is negative] First consider case $e^+ < \overline{e}$. From (13), (4), and (11), $w_nh_n(e^+,s_l) - e^+ = w_l\underline{h}_l$ can be expressed as

$$\left[\alpha A\left(\frac{F(e^+)\underline{h}_l}{\left\{\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right\}\delta\frac{(1-s_l)}{p}}\right)^{1-\alpha}\delta\frac{1-s_l}{p}-1\right]e^+=(1-\alpha)A\left(\frac{\left\{\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right\}\delta\frac{1-s_l}{p}}{F(e^+)\underline{h}_l}\right)^{\alpha}\underline{h}_l$$

$$\Leftrightarrow\left[\alpha A\left(\frac{F(e^+)\underline{h}_l}{\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}}\right)^{1-\alpha}-\left(\delta\frac{1-s_l}{p}\right)^{-\alpha}\right]e^+=(1-\alpha)A\left(\frac{\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}}{F(e^+)\underline{h}_l}\right)^{\alpha}\underline{h}_l \tag{39}$$

Since the LHS increases with e^+, the RHS decreases with e^+, and the LHS decreases with s_l, e^+ satisfying the above equation increases with s_l.

When $e^+=\overline{e}$, from (13), (4), and (12), $w_nh_n(\overline{e},s_l)-\overline{e}=w_l\underline{h}_l$ can be expressed as

$$\left[\alpha A\left(\frac{\{F(\overline{e})+(1-\pi_n)[1-F(\overline{e})]\}\underline{h}_l}{\pi_n[1-F(\overline{e})]\overline{e}}\right)^{1-\alpha}\delta\frac{1-s_l}{p}-1\right]\overline{e}=(1-\alpha)A\left(\frac{\pi_n[1-F(\overline{e})]\overline{e}}{\{F(\overline{e})+(1-\pi_n)[1-F(\overline{e})]\}\underline{h}_l}\right)^{\alpha}\underline{h}_l$$

$$\Leftrightarrow\left[\alpha A\left(\frac{\{F(\overline{e})+(1-\pi_n)[1-F(\overline{e})]\}\underline{h}_l}{\pi_n[1-F(\overline{e})]\overline{e}}\right)^{1-\alpha}-\left(\delta\frac{1-s_l}{p}\right)^{-\alpha}\right]\overline{e}=(1-\alpha)A\left(\frac{\pi_n[1-F(\overline{e})]\overline{e}}{\{F(\overline{e})+(1-\pi_n)[1-F(\overline{e})]\}\underline{h}_l}\right)^{\alpha}\underline{h}_l. \tag{40}$$

Since the LHS decreases with π_n, the RHS increases with π_n, and the LHS decreases with s_l, π_n satisfying the above equation decreases with s_l. ■

Proof of Lemma 2. [When the return to educational investment for local jobs is positive] When $e^+<\overline{e}$, from (2) and (32) in the proof of Lemma 1,

$$\begin{aligned}\frac{\partial w_n}{\partial e^+}&=(1-\alpha)w_n\left[\frac{1}{H_l(e^+,s_l)}\frac{\partial H_l(e^+,s_l)}{\partial e^+}-\frac{1}{H_n(e^+,s_l)}\frac{\partial H_n(e^+,s_l)}{\partial e^+}\right]\\&=(1-\alpha)w_nf(e^+)\left[\frac{h_l(e^+,s_l)}{H_l(e^+,s_l)}+\frac{h_n(e^+,s_l)}{H_n(e^+,s_l)}\right]\\&=w_nf(e^+)\frac{h_n(e^+,s_l)}{H_n(e^+,s_l)}\ \text{(from (10))}\\&=\frac{w_ne^+}{\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}}f(e^+)>0.\end{aligned} \tag{41}$$

From (2), (33), and (34) in the proof of Lemma 1,

$$\begin{aligned}\frac{\partial w_n}{\partial s_l}&=(1-\alpha)w_n\left[\frac{1}{H_l(e^+,s_l)}\frac{\partial H_l(e^+,s_l)}{\partial s_l}-\frac{1}{H_n(e^+,s_l)}\frac{\partial H_n(e^+,s_l)}{\partial s_l}\right]\\&=(1-\alpha)\frac{w_n}{p}\left\{\frac{1}{H_l(e^+,s_l)}\int_0^{e^+}ef(e)de+\frac{\delta}{H_n(e^+,s_l)}\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]\right\}\\&=(1-\alpha)\frac{w_n}{p}\left[\frac{\int_0^{e^+}ef(e)de}{\underline{h}_l\int_0^{e^+}f(e)de+\frac{s_l}{p}\int_0^{e^+}ef(e)de}+\frac{p}{1-s_l}\right]\\&=(1-\alpha)\frac{w_n}{p}\frac{p\underline{h}_l\int_0^{e^+}f(e)de+\int_0^{e^+}ef(e)de}{(1-s_l)\left[\underline{h}_l\int_0^{e^+}f(e)de+\frac{s_l}{p}\int_0^{e^+}ef(e)de\right]}\\&=\frac{w_ne^+}{\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}}\frac{1}{p}\frac{\alpha\left[p\underline{h}_l\int_0^{e^+}f(e)de+\int_0^{e^+}ef(e)de\right]}{(1-s_l)h_l(e^+,s_l)}>0,\end{aligned} \tag{42}$$

where the last equality is because $\alpha\left[\underline{h}_l\int_0^{e^+}f(e)de+\frac{s_l}{p}\int_0^{e^+}ef(e)de\right]e^+=(1-\alpha)\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]h_l(e^+,s_l)$ from (10).

From (41), (42), and (37) in the proof of Lemma 1,

$$\frac{dw_n}{ds_l} = \frac{\partial w_n}{\partial s_l} + \frac{\partial w_n}{\partial e^+}\frac{de^+}{ds_l}$$

$$= \frac{w_n e^+}{\int_{e^+}^{\overline{e}} ef(e)de + [1-F(\overline{e})]\overline{e}}\frac{1}{p}\left(\frac{\alpha\left[p\underline{h}_l F(e^+) + \int_0^{e^+} ef(e)de\right]}{(1-s_l)h_l(e^+,s_l)} + \frac{f(e^+)e^+\left\{(1-\alpha)\left[\int_{e^+}^{\overline{e}} ef(e)de + [1-F(\overline{e})]\overline{e}\right] - \alpha\int_0^{e^+} ef(e)de\right\}}{(1-\alpha)\frac{h_l}{e^+}\left[\int_{e^+}^{\overline{e}} ef(e)de + [1-F(\overline{e})]\overline{e}\right] + h_l(e^+,s_l)e^+ f(e^+)}\right)$$

$$= \frac{w_n e^+}{\int_{e^+}^{\overline{e}} ef(e)de + [1-F(\overline{e})]\overline{e}}\frac{1}{p}\frac{\begin{pmatrix}\alpha\left[p\underline{h}_l F(e^+) + \int_0^{e^+} ef(e)de\right]\left\{(1-\alpha)\frac{h_l}{e^+}\left[\int_{e^+}^{\overline{e}} ef(e)de + [1-F(\overline{e})]\overline{e}\right] + h_l(e^+,s_l)e^+ f(e^+)\right\} \\ +(1-s_l)h_l(e^+,s_l)f(e^+)e^+\left\{(1-\alpha)\left[\int_{e^+}^{\overline{e}} ef(e)de + [1-F(\overline{e})]\overline{e}\right] - \alpha\int_0^{e^+} ef(e)de\right\}\end{pmatrix}}{(1-s_l)h_l(e^+,s_l)\left\{(1-\alpha)\frac{h_l}{e^+}\left[\int_{e^+}^{\overline{e}} ef(e)de + [1-F(\overline{e})]\overline{e}\right] + h_l(e^+,s_l)e^+ f(e^+)\right\}}$$

$$= \frac{w_n e^+}{\int_{e^+}^{\overline{e}} ef(e)de + [1-F(\overline{e})]\overline{e}}\frac{1}{p}\frac{\begin{pmatrix}\alpha\left[p\underline{h}_l F(e^+) + \int_0^{e^+} ef(e)de\right](1-\alpha)\frac{h_l}{e^+}\left[\int_{e^+}^{\overline{e}} ef(e)de + [1-F(\overline{e})]\overline{e}\right] \\ +h_l(e^+,s_l)e^+ f(e^+)\left\{(1-s_l)(1-\alpha)\left[\int_{e^+}^{\overline{e}} ef(e)de + [1-F(\overline{e})]\overline{e}\right] + \alpha p\left[\underline{h}_l F(e^+) + \frac{s_l}{p}\int_0^{e^+} ef(e)de\right]\right\}\end{pmatrix}}{(1-s_l)h_l(e^+,s_l)\left\{(1-\alpha)\frac{h_l}{e^+}\left[\int_{e^+}^{\overline{e}} ef(e)de + [1-F(\overline{e})]\overline{e}\right] + h_l(e^+,s_l)e^+ f(e^+)\right\}}$$

$$= \frac{w_n}{h_l(e^+,s_l)}\frac{e^+}{p}\frac{1}{1-s_l}\frac{(1-\alpha)\left\{\alpha\frac{h_l}{e^+}\left[p\underline{h}_l F(e^+) + \int_0^{e^+} ef(e)de\right] + \left(\frac{p}{e^+}\underline{h}_l + 1\right)h_l(e^+,s_l)e^+ f(e^+)\right\}}{(1-\alpha)\frac{h_l}{e^+}\left[\int_{e^+}^{\overline{e}} ef(e)de + [1-F(\overline{e})]\overline{e}\right] + h_l(e^+,s_l)e^+ f(e^+)} > 0, \tag{43}$$

where the last equality is again from (10).

When $e^+ = \overline{e}$, from (6), (7), and (8),

$$H_l(\pi_n, s_l) = \int_0^{\overline{e}} h_l(e,s_l)f(e)de + \left(1 - \alpha\left\{1 + \frac{\int_0^{\overline{e}} h_l(e,s_l)f(e)de}{[1-F(\overline{e})]h_l(\overline{e},s_l)}\right\}\right)[1 - F(\overline{e})]\, h_l(\overline{e},s_l)$$

$$= (1-\alpha)\left\{\int_0^{\overline{e}} h_l(e,s_l)f(e)de + [1 - F(\overline{e})]\, h_l(\overline{e},s_l)\right\}, \tag{44}$$

$$H_n(\pi_n, s_l) = \alpha\frac{h_n(\overline{e},s_l)}{h_l(\overline{e},s_l)}\left\{\int_0^{\overline{e}} h_l(e,s_l)f(e)de + [1 - F(\overline{e})]\, h_l(\overline{e},s_l)\right\}. \tag{45}$$

By substituting the above equations into (2),

$$w_n = \alpha A\left(\frac{1-\alpha}{\alpha}\frac{h_l(\overline{e},s_l)}{h_n(\overline{e},s_l)}\right)^{1-\alpha}. \tag{46}$$

Thus,

$$\frac{dw_n}{ds_l} = (1-\alpha)w_n\left(\frac{1}{h_l(\overline{e},s_l)}\frac{e}{p} + \frac{1}{h_n(\overline{e},s_l)}\delta\frac{e}{p}\right)$$
$$= (1-\alpha)w_n\frac{1}{1-s_l}\frac{h_l(\overline{e},s_l) + \frac{1}{\delta}h_n(\overline{e},s_l)}{h_l(\overline{e},s_l)} > 0. \tag{47}$$

Since $w_l = (1-\alpha)A\left(\frac{H_n}{H_l}\right)^{\alpha} = (1-\alpha)A\left(\frac{\alpha A}{w_n}\right)^{\frac{\alpha}{1-\alpha}}$ from (2), $\frac{dw_l}{ds_l} = -\frac{\alpha}{1-\alpha}\frac{w_l}{w_n}\frac{dw_n}{ds_l} < 0$.

[When the return is negative] Straightforward from Lemma 1 and the first equation of (39) when $e^+ < \overline{e}$ and of (40) when $e^+ = \overline{e}$. ■

Proof of Proposition 2. When $e^+ < \overline{e}$, $w_l\underline{h}_l = w_n h_n(e^+,s_l) - e^+ = [w_n h_n(1,s_l) - 1]\, e^+$ decreases with s_l from Lemma 2. Then, $w_n h_n(e,s_l) - e = [w_n h_n(1,s_l) - 1]\, e$ for $e > e^+$ too decreases with s_l, because $w_n h_n(1, s_l) - 1$ decreases with s_l from the above equation and $\frac{de^+}{ds_l} > 0$ (Lemma 1). When $e^+ = \bar{e}$, $w_l \underline{h}_l = w_n h_n(\bar{e}, s_l) - \bar{e}$ decreases with s_l from Lemma 2. ■

Proof of Proposition 3. (i) From (47) in the proof of Lemma 2,

$$\begin{aligned}\frac{d[w_n h_n(\bar{e}, s_l)]}{ds_l} &= \frac{dw_n}{ds_l} h_n(\bar{e}, s_l) + w_n \frac{dh_n(\bar{e}, s_l)}{ds_l} \\ &= \frac{w_n h_n(\bar{e},s_l)}{1-s_l} \left[(1-\alpha)\frac{h_l(\bar{e},s_l)+\frac{1}{\delta}h_n(\bar{e},s_l)}{h_l(\bar{e},s_l)} - 1\right] \\ &= \frac{w_n h_n(\bar{e},s_l)}{1-s_l} \frac{(1-\alpha)\frac{1}{\delta}h_n(\bar{e},s_l) - \alpha h_l(\bar{e},s_l)}{h_l(\bar{e},s_l)}. \end{aligned} \tag{48}$$

Thus,

$$\begin{aligned}\frac{d(w_n h_n)}{ds_l} \gtreqless 0 &\Leftrightarrow (1-\alpha)\frac{(1-s_l)\bar{e}}{p} - \alpha\left(\underline{h}_l + \frac{s_l \bar{e}}{p}\right) \gtreqless 0 \\ &\Leftrightarrow s_l \lesseqgtr (1-\alpha) - \alpha \frac{p\underline{h}_l}{\bar{e}}. \end{aligned} \tag{49}$$

(ii) From (47) in the proof of Lemma 2 and (2),

$$\begin{aligned}\frac{d[w_l h_l(e, s_l)]}{ds_l} &= \frac{dw_l}{ds_l} h_l(e, s_l) + w_l \frac{dh_l(e, s_l)}{ds_l} \\ &= -\frac{\alpha}{1-\alpha}\frac{w_l}{w_n}\frac{dw_n}{ds_l} h_l(e, s_l) + w_l \frac{dh_l(e,s_l)}{ds_l} \\ &= \frac{w_l}{1-s_l}\frac{1}{h_l(\bar{e},s_l)}\left\{-\alpha\left[h_l(\bar{e}, s_l) + \tfrac{1}{\delta}h_n(\bar{e}, s_l)\right] h_l(e, s_l) + \tfrac{e}{p}(1-s_l) h_l(\bar{e}, s_l)\right\}. \end{aligned} \tag{50}$$

Thus,

$$\begin{aligned}\frac{d(w_l h_l)}{ds_l} \gtreqless 0 &\Leftrightarrow -\alpha\left(\underline{h}_l + \tfrac{\bar{e}}{p}\right)\left(\underline{h}_l + \tfrac{s_l e}{p}\right) + \tfrac{e}{p}(1-s_l)\left(\underline{h}_l + \tfrac{s_l \bar{e}}{p}\right) \gtreqless 0 \\ &\Leftrightarrow p^2\left\{-e\bar{e}(s_l)^2 + [(1-\alpha)\bar{e} - (1+\alpha)p\underline{h}_l] e s_l + [-\alpha(p\underline{h}_l + \bar{e}) + e] p\underline{h}_l\right\} \gtreqless 0. \end{aligned} \tag{51}$$

(a) Suppose that the LHS of (51) is positive at $s_l = 0$, i.e. $e > \alpha(p\underline{h}_l + \bar{e})$. Because the derivative of the LHS at $s_l = 0$ is non-positive, i.e. $(1-\alpha)\bar{e} - (1+\alpha)p\underline{h}_l \le 0$ and the LHS at $s_l = 1$ equals $-e\bar{e} + [(1-\alpha)\bar{e} - (1+\alpha)p\underline{h}_l] e + [-\alpha(p\underline{h}_l + \bar{e}) + e] p\underline{h}_l = -\alpha(p\underline{h}_l + \bar{e})(p\underline{h}_l + e) < 0$, there exists an $s_l^\star(e) \in (0, 1)$ such that $\frac{d(w_l h_l)}{ds_l} \gtreqless 0 \Leftrightarrow s_l \lesseqgtr s_l^\star(e)$, where $s_l^{\star\prime}(e) > 0$, since, from (51),

$$\begin{aligned} s_l^\star(e) &= \frac{[(1-\alpha)\bar{e} - (1+\alpha)\underline{h}_l] e + \sqrt{[(1-\alpha)\bar{e} - (1+\alpha)p\underline{h}_l]^2 e^2 + 4e\bar{e}[-\alpha(p\underline{h}_l + \bar{e}) + e] p\underline{h}_l}}{2e\bar{e}} \\ &= \frac{[(1-\alpha)\bar{e} - (1+\alpha)\underline{h}_l] + \sqrt{[(1-\alpha)\bar{e} - (1+\alpha)p\underline{h}_l]^2 + 4\bar{e}[-\alpha(p\underline{h}_l + \bar{e}) e^{-1} + 1] p\underline{h}_l}}{2\bar{e}}. \end{aligned} \tag{52}$$

$s_l^\star(e) < s_l^{\star\star}$ for $e < \bar{e}$, since $w_l h_l(\bar{e}, s_l) = w_n h_n(\bar{e}, s_l)$ (thus $s_l^\star(\bar{e}) = s_l^{\star\star}$) and $s_l^{\star\prime}(e) > 0$.

Suppose instead that the LHS of (51) is non-positive at $s_l = 0$, i.e. $e \le \alpha(p\underline{h}_l + \bar{e})$. Because the derivative of the LHS at $s_l = 0$ is non-positive and the LHS at $s_l = 1$ is negative, $\frac{d(w_l h_l)}{ds_l} < 0$ for any $s_l > 0$ (and $\frac{d(w_l h_l)}{ds_l} < (=) 0$ at $s_l = 0$ when $e < (=) \alpha(p\underline{h}_l + \bar{e})$).

(b) The case in which the LHS of (51) at $s_l = 0$ is positive, i.e. $e > \alpha(p\underline{h}_l + \bar{e})$, can be proven as in (a). Suppose that the LHS at $s_l = 0$ is zero, i.e. $e = \alpha\left(p\underline{h}_l + \overline{e}\right)$. Since the derivative of the LHS at $s_l = 0$ is positive, i.e. $(1-\alpha)\frac{\overline{e}}{p} - (1+\alpha)\underline{h}_l > 0$, there exists an $s_l^\star(e)$ in $(0,1)$, such that the LHS of (51) equals 0, and the LHS is zero at $s_l = 0$, positive for $s_l \in (0, s_l^\star(e))$, and negative for $s_l > s_l^\star(e)$. Thus, $\frac{d(w_l h_l)}{ds_l} \gtreqless 0$ for positive $s_l \lesseqgtr s_l^\star(e)$ and $\frac{d(w_l h_l)}{ds_l} = 0$ at $s_l = 0$.

Instead, suppose that the LHS of (51) at $s_l = 0$ is negative, $e < \alpha\left(p\underline{h}_l + \overline{e}\right)$. Since the derivative of the LHS at $s_l = 0$ is positive, i.e. $(1-\alpha)\frac{\overline{e}}{p} - (1+\alpha)\underline{h}_l > 0$, the LHS is positive (negative) when $e > (<) \frac{\alpha(p\underline{h}_l+\overline{e})p\underline{h}_l}{\{-\overline{e}s_l + [(1-\alpha)\overline{e} - (1+\alpha)p\underline{h}_l]\}s_l + p\underline{h}_l}$, where $\frac{\alpha(p\underline{h}_l+\overline{e})p\underline{h}_l}{\{-\overline{e}s_l + [(1-\alpha)\overline{e} - (1+\alpha)p\underline{h}_l]\}s_l + p\underline{h}_l} < \alpha\left(p\underline{h}_l + \overline{e}\right)$ when $s_l > 0$ and $-\overline{e}s_l + [(1-\alpha)\overline{e} - (1+\alpha)p\underline{h}_l] > 0 \Leftrightarrow s_l \in (0, \left[(1-\alpha) - (1+\alpha)\frac{p\underline{h}_l}{\overline{e}}\right])$.

So the LHS of (51) is negative for any $e < \alpha\left(p\underline{h}_l + \overline{e}\right)$ when $s_l \geq (1-\alpha) - (1+\alpha)\frac{p\underline{h}_l}{\overline{e}}$. For $s_l < (1-\alpha) - (1+\alpha)\frac{p\underline{h}_l}{\overline{e}}$, the LHS of (51) is positive (negative) when $e > (<) \frac{\alpha(p\underline{h}_l+\overline{e})p\underline{h}_l}{\{-\overline{e}s_l + [(1-\alpha)\overline{e} - (1+\alpha)p\underline{h}_l]\}s_l + p\underline{h}_l}$, where the RHS is lowest when $-\overline{e}s_l + [(1-\alpha)\overline{e} - (1+\alpha)p\underline{h}_l] - \overline{e}s_l = 0 \Leftrightarrow s_l = \frac{(1-\alpha)\overline{e} - (1+\alpha)p\underline{h}_l}{2\overline{e}}$.

Hence, the LHS of (51) is negative, i.e. $\frac{d(w_l h_l)}{ds_l} < 0$, for any s_l when

$$e \leq \left[\frac{\alpha(p\underline{h}_l+\overline{e})p\underline{h}_l}{\{-\overline{e}s_l + [(1-\alpha)\overline{e} - (1+\alpha)p\underline{h}_l]\}s_l + p\underline{h}_l} \text{ at } s_l = \frac{(1-\alpha)\overline{e} - (1+\alpha)p\underline{h}_l}{2\overline{e}}\right] = \frac{\alpha(p\underline{h}_l+\overline{e})p\underline{h}_l}{p\underline{h}_l + \frac{[(1-\alpha)\overline{e} - (1+\alpha)p\underline{h}_l]^2}{4\overline{e}}} = \frac{\alpha(p\underline{h}_l+\overline{e})}{1 + \frac{1}{4p\underline{h}\overline{e}}[(1-\alpha)\overline{e} - (1+\alpha)p\underline{h}_l]^2},$$

except at $s_l = \frac{(1-\alpha)\overline{e} - (1+\alpha)p\underline{h}_l}{2\overline{e}}$ and $e = \frac{\alpha(p\underline{h}_l+\overline{e})}{1 + \frac{1}{4p\underline{h}\overline{e}}[(1-\alpha)\overline{e} - (1+\alpha)p\underline{h}_l]^2}$, in which the derivative is zero.

When $e \in \left(\frac{\alpha(p\underline{h}_l+\overline{e})}{1 + \frac{1}{4p\underline{h}\overline{e}}[(1-\alpha)\overline{e} - (1+\alpha)p\underline{h}_l]^2}, \alpha\left(p\underline{h}_l + \overline{e}\right)\right)$, there exist $s_l^\circ(e)$ and $s_l^\star(e)$, where $0 < s_l^\circ(e) < s_l^\star(e) < (1-\alpha) - (1+\alpha)\frac{p\underline{h}_l}{\overline{e}} < s_l^{\star\star}$, such that the LHS of (51) equals 0 ($s_l^\circ(e)$ is the smaller solution), and the LHS is negative for $s_l < s_l^\circ(e)$, positive for $s_l \in (s_l^\circ(e), s_l^\star(e))$, and negative for $s_l > s_l^\star(e)$. $s_l^{\star\prime}(e) > 0 > s_l^{\circ\prime}(e)$ from (52). ■

Proof of Proposition 5. $e^+ = \overline{e}$ always holds when $F(\overline{e}) \leq 1-\alpha$ from Proposition 1 (ii). The following lemma is used in the proof of the proposition.

Lemma A2 *When $e^+ = \overline{e}$, net earnings change continuously when the return to educational investment for local jobs turns from negative to positive with a change in s_l.*

Proof of Lemma A2. When $e^+ = \overline{e}$ and the return is negative, $w_n h_n(\overline{e}, s_l) - \overline{e} = w_l h_l(0, s_l) \Leftrightarrow \left[\alpha A\left(\frac{H_l}{H_n}\right)^{1-\alpha}\delta^{\frac{1-s_l}{p}} - 1\right]\overline{e} = (1-\alpha)A\left(\frac{H_n}{H_l}\right)^{\alpha}\underline{h}_l$, where $H_l = \{F(\overline{e}) + (1-\pi_n)[1-F(\overline{e})]\}\,\underline{h}_l$ and $H_n = \pi_n[1-F(\overline{e})]h_n(\overline{e}, s_l)$, from (12) and (13). When $e^+ = \overline{e}$ and the return is positive, $w_n h_n(\overline{e}, s_l) - \overline{e} = w_l h_l(\overline{e}, s_l) - \overline{e} \Leftrightarrow \left[\alpha A\left(\frac{H_l}{H_n}\right)^{1-\alpha}\delta^{\frac{1-s_l}{p}} - 1\right]\overline{e} = (1-\alpha)A\left(\frac{H_n}{H_l}\right)^{\alpha}\left(\underline{h}_l + \frac{s_l}{p}\overline{e}\right) - \overline{e}$, where $H_l = \int_0^{\overline{e}} h_l(e, s_l)f(e)de + (1-\pi_n)[1-F(\overline{e})]h_l(\overline{e}, s_l)$ and $H_n = \pi_n[1-F(\overline{e})]h_n(\overline{e}, s_l)$, from (2), (6), and (7). When the return is zero, i.e. $w_l\frac{s_l}{p} - 1 = 0$, this equation becomes $\left[\alpha A\left(\frac{H_l}{H_n}\right)^{1-\alpha}\delta^{\frac{1-s_l}{p}} - 1\right]\overline{e} = (1-\alpha)A\left(\frac{H_n}{H_l}\right)^{\alpha}\underline{h}_l$, the same equation as the case of the negative return, and net earnings of all workers are the same. Since $\frac{H_n}{H_l}$ satisfying this equation is uniquely determined for given s_l, the net earnings when the return is zero are same as the net earnings when the return is negative and $s_l \to \underline{s_l}, \overline{s_l}$. ■

(i) Earnings decrease with s_l when the return to educational investment for local jobs is negative from Proposition 2, while when the return is positive and $e^+ = \overline{e}$, earnings of all decrease with s_l for $s_l > s_l^{\star\star} \equiv (1-\alpha) - \alpha\frac{p\underline{h}_l}{\overline{e}}$ from Proposition 3. From Proposition 1 (i)(b), the lower critical value of s_l for the negative return, $\underline{s_l}$, decreases with A and δ and increases with p. Hence, from

Lemma A2, net earnings of all decrease with s_l when A and δ are small enough or p is large enough that $\underline{s_l} \geq s_l^{\star\star} = s_l^\star(e^+) = s_l^\star(\bar{e})$. Note that $s_l^{\star\star}$ is smaller than the higher critical value $\overline{s_l}$, because $s_l^{\star\star} < 1-\alpha < \overline{s_l}$ from Lemma A1 in the proof of Proposition 1.

(ii) (a) When $\underline{s_l} < s_l^{\star\star} = s_l^\star(e^+) = s_l^\star(\bar{e})$, from Propositions 2 and 3 and Lemma A2, net earnings of workers with a national job decrease with s_l for $s_l < \underline{s_l}$, increase with s_l for $s_l \in (\underline{s_l}, s_l^{\star\star})$, and decrease with s_l for $s_l > s_l^{\star\star}$.

Their net earnings are maximized at either $s_l = 0$ or $s_l = s_l^{\star\star}$. From the proof of Lemma A2, net earnings of such workers at $s_l = 0$ equal

$$\left[\alpha A\left(\frac{\{F(\bar{e})+(1-\pi_n)[1-F(\bar{e})]\}\underline{h}_l}{\pi_n[1-F(\bar{e})]\frac{\delta}{p}\bar{e}}\right)^{1-\alpha}\frac{\delta}{p}-1\right]\bar{e}, \tag{53}$$

where π_n is a solution to (28) in the proof of Proposition 1.

From (23) in the proof of Proposition 1 and (2), net earnings of such workers at $s_l = s_l^{\star\star}$ equal

$$\left\{\alpha A\left[\frac{(1-\alpha)\left(\underline{h}_l+\frac{s_l^{\star\star}\bar{e}}{p}\right)}{\alpha\delta\frac{(1-s_l^{\star\star})\bar{e}}{p}}\right]^{1-\alpha}\delta\frac{1-s_l^{\star\star}}{p}-1\right\}\bar{e} = \left\{\alpha A\left[\frac{(1-\alpha)^2}{\alpha^2\delta}\right]^{1-\alpha}\frac{\delta}{p}\alpha\left(1+\frac{p\underline{h}_l}{\bar{e}}\right)-1\right\}\bar{e}. \tag{54}$$

From these equations, the net earnings are maximized at $s_l = s_l^{\star\star}$ if

$$\left[\frac{(1-\alpha)^2}{\alpha^2}\right]^{1-\alpha}\alpha\left(1+\frac{p\underline{h}_l}{\bar{e}}\right) > \left(\frac{\{F(\bar{e})+(1-\pi_n)[1-F(\bar{e})]\}\underline{h}_l}{\pi_n[1-F(\bar{e})]\frac{\bar{e}}{p}}\right)^{1-\alpha}, \tag{55}$$

which holds when A and δ are large and p is small, since π_n increases with A and δ and decreases with p, which can be shown from (28), $\left(1+\frac{p\underline{h}_l}{\bar{e}}\right)\left(\frac{\bar{e}}{p}\right)^{1-\alpha}$ decreases with p from $-(1-\alpha)\frac{1}{p} + \left(1+\frac{p\underline{h}_l}{\bar{e}}\right)^{-1}\frac{\underline{h}_l}{\bar{e}} = \frac{-(1-\alpha)+\alpha\frac{p\underline{h}_l}{\bar{e}}}{p\left(1+\frac{p\underline{h}_l}{\bar{e}}\right)} = \frac{-s_l^{\star\star}}{p\left(1+\frac{p\underline{h}_l}{\bar{e}}\right)} < 0$, and the condition does hold when $\pi_n = 1$, i.e. $\left[\frac{(1-\alpha)^2}{\alpha^2}\right]^{1-\alpha}\alpha\left(1+\frac{p\underline{h}_l}{\bar{e}}\right) > \left(\frac{F(\bar{e})\underline{h}_l}{[1-F(\bar{e})]\frac{\bar{e}}{p}}\right)^{1-\alpha}$. Since $F(\bar{e}) \leq 1-\alpha$ at $s_l = 0$ from (27) in the proof of Proposition 1, the last statement is proved if $\left[\frac{(1-\alpha)^2}{\alpha^2}\right]^{1-\alpha}\alpha\left(1+\frac{p\underline{h}_l}{\bar{e}}\right) > \left(\frac{(1-\alpha)\underline{h}_l}{\alpha\frac{\bar{e}}{p}}\right)^{1-\alpha} \Leftrightarrow (1-\alpha)^{1-\alpha}\alpha^\alpha\left(1+\left(\frac{p\underline{h}_l}{\bar{e}}\right)^{-1}\right)\left(\frac{p\underline{h}_l}{\bar{e}}\right)^\alpha > 1$ holds, which is true, because $s_l^{\star\star} \equiv (1-\alpha)-\alpha\frac{p\underline{h}_l}{\bar{e}} > 0 \Leftrightarrow \frac{p\underline{h}_l}{\bar{e}} < \frac{1-\alpha}{\alpha}$ and the LHS decreases with $\frac{p\underline{h}_l}{\bar{e}}$:

$$\frac{\alpha}{\frac{p\underline{h}_l}{\bar{e}}} - \frac{\left(\frac{p\underline{h}_l}{\bar{e}}\right)^{-2}}{1+\left(\frac{p\underline{h}_l}{\bar{e}}\right)^{-1}} = \frac{\alpha-(1-\alpha)\left(\frac{p\underline{h}_l}{\bar{e}}\right)^{-1}}{\frac{p\underline{h}_l}{\bar{e}}\left[1+\left(\frac{p\underline{h}_l}{\bar{e}}\right)^{-1}\right]} < 0 \text{ from } \frac{p\underline{h}_l}{\bar{e}} < \frac{1-\alpha}{\alpha}. \tag{56}$$

(b) Consider case $\bar{e} \leq \frac{1+\alpha}{1-\alpha}p\underline{h}_l$ first. Since $s_l^\star(e)$ increases with e, $s_l^\star(\alpha(p\underline{h}_l+\bar{e})) = 0$, and $s_l^\star(\bar{e}) = s_l^{\star\star}$ from (52) in the proof of Proposition 3, there exists an $e^\dagger \in (\alpha(p\underline{h}_l+\bar{e}), \bar{e})$, such that $s_l^\star(e^\dagger) = \underline{s_l}$. Then, the relationship between s_l and net earnings of workers with a local job and $e > e^\dagger$ is similar to that of workers with a national job: their net earnings decrease with s_l for

$s_l < \underline{s_l}$, increase with s_l for $s_l \in (\underline{s_l}, s_l^\star(e))$, and decrease with s_l for $s_l > s_l^\star(e)$ from Propositions 2 and 3 and Lemma A2. As for workers with $e \leq e^\dagger$, net earnings decrease with s_l.

Now consider case $\overline{e} > \frac{1+\alpha}{1-\alpha} p\underline{h}_l$. If A and δ are small enough or p is large enough that $s_l^\star(\alpha(p\underline{h}_l+\overline{e})) = 1-\alpha - \frac{(1+\alpha)p\underline{h}_l}{\overline{e}} < \underline{s_l}$, the result is same as the corresponding case of $\overline{e} \leq \frac{1+\alpha}{1-\alpha} p\underline{h}_l$.

Let $\Lambda(\overline{e}) \equiv \frac{\alpha(p\underline{h}_l+\overline{e})}{1+\frac{1}{4p\underline{h}\overline{e}}[(1-\alpha)\overline{e}-(1+\alpha)p\underline{h}_l]^2}$. If $s_l^\star(\alpha(p\underline{h}_l+\overline{e})) > \underline{s_l} > s_l^\star(\Lambda(\overline{e})) = \frac{(1-\alpha)-\frac{(1+\alpha)p\underline{h}_l}{\overline{e}}}{2}$, there exists an $e^\ddagger \in (\Lambda(\overline{e}), \alpha(p\underline{h}_l+\overline{e}))$ such that $s_l^\star(e^\ddagger) = \underline{s_l}$. The results for workers with $e > \alpha(p\underline{h}_l+\overline{e})$ and workers with $e \leq \Lambda(\overline{e})$ are same as the corresponding cases ($e > e^\dagger$ and $e \leq e^\dagger$ respectively) of $\overline{e} \leq \frac{1+\alpha}{1-\alpha} p\underline{h}_l$. As for workers with $e \in (\Lambda(\overline{e}), \alpha(p\underline{h}_l+\overline{e}))$, $s_l^\circ(e) \leq \frac{(1-\alpha)-\frac{(1+\alpha)p\underline{h}_l}{\overline{e}}}{2} = s_l^\star(\Lambda(\overline{e})) < s_l^\star(e^\ddagger) = \underline{s_l}$ holds for any e from (51) and (52) in the proof of Proposition 3. Hence, the result of workers with $e > e^\ddagger$ is similar to that of workers with $e > \alpha(p\underline{h}_l+\overline{e})$, and the result of workers with $e \leq e^\ddagger$ is same as that of workers with $e \leq \Lambda(\overline{e})$. To summarize, the result is similar to the corresponding case of $\overline{e} \leq \frac{1+\alpha}{1-\alpha} p\underline{h}_l$ except that the critical wealth level is $e^\ddagger$, not $e^\dagger$.

Finally, if $s_l^\star(\Lambda(\overline{e})) > \underline{s_l}$, because $s_l^\circ(e)$ decreases with e, $s_l^-(\alpha(p\underline{h}_l+\overline{e})) = 0$, and $s_l^-(\Lambda(\overline{e})) = s_l^\star(\Lambda(\overline{e})) = \frac{(1-\alpha)-\frac{(1+\alpha)p\underline{h}_l}{\overline{e}}}{2}$ from (51) and (52), there exists an $e^* \in (\Lambda(\overline{e}), \alpha(p\underline{h}_l+\overline{e}))$ such that $s_l^-(e^*) = \underline{s_l}$. Hence, as for workers with $e \in (\Lambda(\overline{e}), \alpha(p\underline{h}_l+\overline{e}))$, net earnings of workers with $e < e^*$ decrease with s_l for $s_l < s_l^\circ(e)$, increase with s_l for $s_l < (s_l^\circ(e), s_l^\star(e))$, and decrease with s_l for $s_l > s_l^\star(e)$, while net earnings of workers with $e \geq e^*$ decrease with s_l for $s_l < \underline{s_l}$, increase with s_l for $s_l \in (\underline{s_l}, s_l^\star(e))$, and decrease with s_l for greater s_l. The results for workers with $e > \alpha(p\underline{h}_l+\overline{e})$ and for workers with $e \leq \Lambda(\overline{e})$ are same as the previous case.

To summarize, net earnings of workers with wealth and thus educational spending greater than a certain level decrease with s_l for $s_l < \max\{\underline{s_l}, s_l^\circ(e)\}$, increase with s_l for $s_l \in (\max\{\underline{s_l}, s_l^\circ(e)\}, s_l^\star(e))$, and decrease with s_l for $s_l > s_l^\star(e)$, while net earnings of workers with $a = e$ smaller than the threshold decrease with s_l for any s_l.

Thus, net earnings of workers with wealth greater than a threshold are maximized at either $s_l = 0$ or $s_l = s_l^\star(e)$. From (28) in the proof of Proposition 1, net earnings of such workers at $s_l = 0$ is same as net earnings of workers with a national job, which equals (53).

From (23) in the proof of Proposition 1, net earnings of such workers with e at $s_l = s_l^\star(e)$ equal

$$(1-\alpha)A\left[\frac{\alpha\delta\frac{(1-s_l^\star(e))\overline{e}}{p}}{(1-\alpha)\left(\underline{h}_l+\frac{s_l^\star(e)\overline{e}}{p}\right)}\right]^\alpha \left(\underline{h}_l+\frac{s_l^\star(e)\overline{e}}{p}\right)-e. \tag{57}$$

Thus, the net earnings are maximized at $s_l = s_l^\star(e)$ if

$$A\left(\delta\frac{\overline{e}}{p}\right)^\alpha\left\{(1-\alpha)\left[\frac{\alpha(1-s_l^\star(e))}{1-\alpha}\right]^\alpha\left(\underline{h}_l+\frac{s_l^\star(e)\overline{e}}{p}\right)^{1-\alpha}-\alpha\left(\frac{\{F(\overline{e})+(1-\pi_n)[1-F(\overline{e})]\}\underline{h}_l}{\pi_n[1-F(\overline{e})]}\right)^{1-\alpha}\right\}+\overline{e}-e>0, \tag{58}$$

which holds when A and δ are large or p is small, because π_n increases with A and δ and decreases with p, the derivative of the first term of the expression inside the large curcly bracket with resepct to s_l is 0 at $s_l = s_l^\star(e)$ and negative for $s_l > s_l^\star(e)$, and thus the expression inside the curly bracket is positive when A and δ are large or p is small from (55):

$$(1-\alpha)\left[\tfrac{\alpha(1-s_l^\star(e))}{1-\alpha}\right]^\alpha\left(\underline{h}_l+\tfrac{s_l^\star(e)\overline{e}}{p}\right)^{1-\alpha}-\alpha\left(\tfrac{\{F(\overline{e})+(1-\pi_n)[1-F(\overline{e})]\}\underline{h}_l}{\pi_n[1-F(\overline{e})]}\right)^{1-\alpha}$$
$$>(1-\alpha)\left[\tfrac{\alpha(1-s_l^{\star\star})}{1-\alpha}\right]^\alpha\left(\underline{h}_l+\tfrac{s_l^{\star\star}\overline{e}}{p}\right)^{1-\alpha}-\alpha\left(\tfrac{\{F(\overline{e})+(1-\pi_n)[1-F(\overline{e})]\}\underline{h}_l}{\pi_n[1-F(\overline{e})]}\right)^{1-\alpha}\quad\text{(from Proposition 3 (ii)(a))}$$
$$=\alpha\left(\tfrac{\overline{e}}{p}\right)^{1-\alpha}\left\{\left[\tfrac{(1-\alpha)^2}{\alpha^2}\right]^{1-\alpha}\alpha\left(1+\tfrac{p\underline{h}_l}{\overline{e}}\right)-\left(\tfrac{\{F(\overline{e})+(1-\pi_n)[1-F(\overline{e})]\}\underline{h}_l}{\pi_n[1-F(\overline{e})]\frac{\overline{e}}{p}}\right)^{1-\alpha}\right\}>0. \tag{59}$$

(c) From the proof of (b), the threshold wealth when $\overline{e}\le\frac{1+\alpha}{1-\alpha}p\underline{h}_l$ is $e^\dagger\in(\alpha(p\underline{h}_l+\overline{e}),\overline{e})$ such that $s_l^\star(e^\dagger)=\underline{s_l}$. The threshold when $\overline{e}>\frac{1+\alpha}{1-\alpha}p\underline{h}_l$ is $e^\dagger\in(\alpha(p\underline{h}_l+\overline{e}),\overline{e})$ if $s_l^\star(\alpha(p\underline{h}_l+\overline{e}))<\underline{s_l}$, $e^\ddagger\in(\Lambda(\overline{e}),\alpha(p\underline{h}_l+\overline{e}))$ such that $s_l^\star(e^\ddagger)=\underline{s_l}$ if $s_l^\star(\alpha(p\underline{h}_l+\overline{e}))>\underline{s_l}>s_l^\star(\Lambda(\overline{e}))$, and $\Lambda(\overline{e})$ if $s_l^\star(\Lambda(\overline{e}))>\underline{s_l}$.

When $\overline{e}>\frac{1+\alpha}{1-\alpha}p\underline{h}_l$, case $s_l^\star(\alpha(p\underline{h}_l+\overline{e}))<\underline{s_l}$ is realized when A and δ are small and p is large, and case $s_l^\star(\Lambda(\overline{e}))>\underline{s_l}$ is realized when A and δ are large and p is small, because $\underline{s_l}$ decreases with A and δ and increases with p from Proposition 1, and $s_l^\star(\alpha(p\underline{h}_l+\overline{e}))=(1-\alpha)-(1+\alpha)\frac{p\underline{h}_l}{\overline{e}}$ and $s_l^\star(\Lambda(\overline{e}))=\frac{1}{2}\left[(1-\alpha)-(1+\alpha)\frac{p\underline{h}_l}{\overline{e}}\right]$ decrease with p. Further, $e^\dagger$ and $e^\ddagger$ decrease with A and δ and increase with p, because $s_l^\star(e)$ decreases with p at $e=e^\dagger$, $e^\ddagger$ from Lemma A3 below and increases with e, and $\underline{s_l}$ decreases with A and δ and increases with p. Hence, the threshold of wealth decreases with A and δ (except when the threshold is $\alpha(p\underline{h}_l+\overline{e})$ and $\Lambda(\overline{e})$) and increases with p. ■

Lemma A3 $\frac{\partial s_l^\star(e^\dagger)}{\partial p}<0$ *and* $\frac{\partial s_l^\star(e^\ddagger)}{\partial p}<0$

Proof of Lemma A3. From (52) in the proof of Proposition 2, the derivative of $s_l^\star(e)$ with respect to p equals a constant times

$$-(1+\alpha)\underline{h}_l+\frac{1}{2}\left\{-(1+\alpha)\underline{h}_l2\left[(1-\alpha)\overline{e}-(1+\alpha)p\underline{h}_l\right]+4\overline{e}\left[-\alpha(p\underline{h}_l+\overline{e})(e)^{-1}+1\right]\underline{h}_l-4\overline{e}\alpha p\underline{h}_l(e)^{-1}\underline{h}_l\right\}$$
$$\times\left\{\left[(1-\alpha)\overline{e}-(1+\alpha)p\underline{h}_l\right]^2+4p\overline{e}\left[-\alpha(p\underline{h}_l+\overline{e})(e)^{-1}+1\right]\underline{h}_l\right\}^{-1/2}. \tag{60}$$

Thus,

$$\frac{\partial s_l^\star(e)}{\partial p}\lesseqgtr 0\Leftrightarrow-(1+\alpha)\underline{h}_l2\left[(1-\alpha)\overline{e}-(1+\alpha)p\underline{h}_l\right]+4\overline{e}\left[-\alpha(p\underline{h}_l+\overline{e})(e)^{-1}+1\right]\underline{h}_l-4\overline{e}\alpha p\underline{h}_l(e)^{-1}\underline{h}_l\lesseqgtr 2(1+\alpha)\underline{h}_l$$
$$\times\left\{\left[(1-\alpha)\overline{e}-(1+\alpha)p\underline{h}_l\right]^2+4p\overline{e}\left[-\alpha(p\underline{h}_l+\overline{e})(e)^{-1}+1\right]\underline{h}_l\right\}^{1/2}. \tag{61}$$

The lemma is proved if it is shown that $\frac{\partial s_l^\star(e)}{\partial p}\ge 0$ cannot hold at $e=e^\dagger,e^\ddagger$. From the above equation, $\frac{\partial s_l^\star(e)}{\partial p}\ge 0$ is possible only when the LHS of the equation is positive, which is true only when $(1-\alpha)\overline{e}-(1+\alpha)p\underline{h}_l<0\Leftrightarrow\overline{e}<\frac{1+\alpha}{1-\alpha}p\underline{h}_l$ or $-\alpha(p\underline{h}_l+\overline{e})(e)^{-1}+1-\alpha p\underline{h}_l(e)^{-1}>0\Leftrightarrow e>\alpha(2p\underline{h}_l+\overline{e})$. When the LHS is positive, the above equation can be expressed as

$$\frac{\partial s_l^\star(e)}{\partial p}\lesseqgtr 0\Leftrightarrow\left\{-(1+\alpha)\underline{h}_l2\left[(1-\alpha)\overline{e}-(1+\alpha)p\underline{h}_l\right]+4\overline{e}\left[-\alpha(p\underline{h}_l+\overline{e})(e)^{-1}+1\right]\underline{h}_l-4\overline{e}\alpha p\underline{h}_l(e)^{-1}\underline{h}_l\right\}^2$$
$$\lesseqgtr\left[2(1+\alpha)\underline{h}_l\right]^2\left\{\left[(1-\alpha)\overline{e}-(1+\alpha)p\underline{h}_l\right]^2+4p\overline{e}\left[-\alpha(p\underline{h}_l+\overline{e})(e)^{-1}+1\right]\underline{h}_l\right\}$$
$$\Leftrightarrow-(1+\alpha)\left[(1-\alpha)\overline{e}-(1+\alpha)p\underline{h}_l\right]\left[-\alpha(2p\underline{h}_l+\overline{e})(e)^{-1}+1\right]+\left[-\alpha(2p\underline{h}_l+\overline{e})(e)^{-1}+1\right]^2\overline{e}$$

$$\lesseqgtr (1+\alpha)^2 p\Big[-\alpha(p\underline{h}_l+\overline{e})(e)^{-1}+1\Big]\,\underline{h}_l$$

$$\Leftrightarrow \Big[-\alpha(2p\underline{h}_l+\overline{e})(e)^{-1}+1\Big]\Big[-(2p\underline{h}_l+\overline{e})(e)^{-1}+\alpha\Big]\,\overline{e}-(1+\alpha)^2p(p\underline{h}_l)\,(e)^{-1}\,\underline{h}_l \lesseqgtr 0. \tag{62}$$

The expression is clearly negative when $e \geq \alpha(2p\underline{h}_l+\overline{e})$. Hence, $\frac{\partial s_l^{\star}(e)}{\partial p} \geq 0$ is possible only when $\overline{e} < \frac{1+\alpha}{1-\alpha}p\underline{h}_l$ and $e < \alpha(2p\underline{h}_l+\overline{e})$. In this case, the LHS of (61) is weakly smaller than $(1+\alpha)\underline{h}_l 2\left[(1+\alpha)p\underline{h}_l-(1-\alpha)\overline{e}\right]$, while, since $e^{\dagger} > \alpha(p\underline{h}_l+\overline{e})$($e^{\ddagger}$ does not exist) when $\overline{e} < \frac{1+\alpha}{1-\alpha}p\underline{h}_l$, the RHS of (61) is greater than $(1+\alpha)\underline{h}_l 2\left[(1+\alpha)p\underline{h}_l-(1-\alpha)\overline{e}\right]$. Hence, $\frac{\partial s_l^{\star}(e^{\dagger})}{\partial p} < 0$ holds in this case. The fact that $e^{\ddagger}$ does not exist when $\overline{e} < \frac{1+\alpha}{1-\alpha}p\underline{h}_l$ proves that $\frac{\partial s_l^{\star}(e^{\ddagger})}{\partial p} \geq 0$ cannot happen. ■

Proof of Corollary 1. Since $F(\overline{e}) = 0 < 1-\alpha$, the proof of Proposition 5 can be applied with $a > \overline{e}$ for all individuals. The result is straightforward from the proof, since $e = \overline{e} = e^{+}$ and thus $s_l^{\star}(e) = s_l^{\star}(\overline{e}) = s_l^{\star\star}$ hold for workers choosing a local job, and $\overline{e} > \alpha(p\underline{h}_l+\overline{e})$ holds by assumption (see footnote 23). ■

Proof of Proposition 6. In order to prove the proposition, the following lemma is used.

Lemma A4 *When $e^{+} < \overline{e}$, net earnings of workers with given wealth and a national job increase discontinuously and those of workers with a local job decrease discontinuously, when the return to educational investment for local jobs turns from negative to positive with a change in s_l.*

Proof of Lemma A4. When $e^{+} < \overline{e}$ and the return is negative, $w_n h_n(e^{+}, s_l) - e^{+} = w_l\underline{h}_l \Leftrightarrow \left[\alpha A\left(\frac{H_l}{H_n}\right)^{1-\alpha}\delta\frac{1-s_l}{p}-1\right]e^{+} = (1-\alpha)A\left(\frac{H_n}{H_l}\right)^{\alpha}\underline{h}_l$, where $H_l = F(e^{+})\underline{h}_l$ and $H_n = \int_{e^{+}}^{\overline{e}} h_n(e, s_l)f(e)de + [1-F(\overline{e})]h_n(\overline{e}, s_l)$, holds from (11) and (13). When $e^{+} < \overline{e}$ and the return is positive, $w_n h_n(e^{+}, s_l) - e^{+} = w_l h_l(e^{+}, s_l) - e^{+} \Leftrightarrow \left[\alpha A\left(\frac{H_l}{H_n}\right)^{1-\alpha}\delta\frac{1-s_l}{p}-1\right]e^{+} = (1-\alpha)A\left(\frac{H_n}{H_l}\right)^{\alpha}\left(\underline{h}_l+\frac{s_l}{p}e^{+}\right) - e^{+}$, where $H_l = \int_0^{e^{+}} h_l(e, s_l)f(e)de$ and $H_n = \int_{e^{+}}^{\overline{e}} h_n(e, s_l)f(e)de + [1-F(\overline{e})]h_n(\overline{e}, s_l)$, holds from (2) and (5). When the return is zero, i.e. $w_l\frac{s_l}{p}-1 = 0$, this equation becomes $\left[\alpha A\left(\frac{H_l}{H_n}\right)^{1-\alpha}\delta\frac{1-s_l}{p}-1\right]e^{+} = (1-\alpha)A\left(\frac{H_n}{H_l}\right)^{\alpha}\underline{h}_l$, the same equation as the case of the negative return. Because $\frac{H_n}{H_l}$ under the negative return is greater than $\frac{H_n}{H_l}$ under the positive return for given e^{+} and $\frac{H_n}{H_l}$ decreases with e^{+}, e^{+} and $\frac{H_n}{H_l}$ satisfying the above equation are greater when the return is negative. Hence, net earnings of workers with a local job are greater and net earnings of workers with given $a(> e^{+})$ and a national job are smaller when the return is negative and s_l approaches a value at which the return is zero than when the return is zero. That is, net earnings of workers with a national job increase discontinuously and those of workers with a local job decrease discontinuously when the return turns from negative to positive with a change in s_l. ■

(i) Let the lower (higher) s_l such that the return to educational investment for local jobs is 0 when $e^{+} < \overline{e}$ be $\underline{s_l}(f)$ ($\overline{s_l}(f)$), whose existence is proved in the proof of Proposition 1. As for workers who have abundant wealth and thus choose a national job for any s_l, i.e. $a \geq \min\{e^{+}(1), \overline{e}\}$, their net earnings decrease with s_l except at $s_l = \underline{s_l}(f)$, where the earnings increase discontinuously from Lemma A4, if $e^{+}(0) \leq \frac{\alpha p\underline{h}_l}{1-\alpha}$ or $\underline{s_l}(f) \geq s_l^{\sharp}$ satisfying $s_l^{\sharp} = (1-\alpha) - \alpha\frac{p\underline{h}_l}{e^{+}(s_l^{\sharp})}$ from Proposition

A1 (i). The former condition holds when p is large, because $e^+(0)$ does not depend on A, δ, and p from (C3) in the proof of Proposition A1 in Appendix C. The latter condition holds when p is large or A and δ are small, because $\underline{s_l}(f)$ increases with p and decreases with A and δ from the proof of Proposition 1, particularly (14), and $s_l^\sharp$ decreases with p.

As for workers who have limited wealth and thus choose a local job for any s_l, i.e. $a = e < e^+(0)$, their net earnings decrease with s_l except at $s_l = \overline{s_l}(f)$, where the earnings increase discontinuously from Lemma A4, when $E(e|e < e^+(0)) \leq \frac{1+\alpha}{1-\alpha}p\underline{h}_l$ and $e \leq \alpha(p\underline{h}_l + E(e|e<e^+(0)))$, when $E(e|e < e^+(0)) > \frac{1+\alpha}{1-\alpha}p\underline{h}_l$ and $e \leq \Omega(e^+(0))$, or when $\underline{s_l}(f) \geq s^{\triangledown}_{l,h}(e)$ for greater e, where $s^{\triangledown}_{l,h}(e)$ is the greater solution of $M(s_l) = 0$ ($M(s_l)$ equals $L(s_l)$ with $e^+(s_l)$ replaced with $E(e|e<e^+(0)) \equiv \frac{\int_0^{e^+(s_l)} ef(e)de}{F(e^+(s_l))}$), from Proposition A1 (ii). Thus, irrespective of $a = e$, their net earnings decrease with s_l except at $s_l = \overline{s_l}(f)$, if $\underline{s_l}(f) \geq s^{\triangledown}_{l,h}(e)$ is true. The condition holds when A and δ are small, because $\underline{s_l}(f)$ decreases with A and δ, while $s^{\triangledown}_{l,h}(e)$ does not depend on A and δ from (C9) in the proof of Proposition A1 in Appendix C and the definition of $s^{\triangledown}_{l,h}(e)$. The condition holds when p is large enough that $\underline{s_l}(f) \geq 1 - \alpha \geq s^{\triangledown}_{l,h}(e)$ holds, where $\underline{s_l}(f) \geq 1 - \alpha$ is from $\underline{s_l}(f)$ being increasing in p, and from (14) and Lemma A1 in the proof of Proposition 1, while $1 - \alpha \geq s^{\triangledown}_{l,h}(e)$ is from $M(1-\alpha) = -\alpha[(p\underline{h}_l + E(e|e<e^+(0))) - \alpha e]p\underline{h}_l \leq 0$, which is true for any e and s_l when p is large enough that $(p\underline{h}_l + E(e|e<e^+(0))) - \alpha e^+(0) \geq 0$.

Finally, as for workers who choose a national (local) job when s_l is small (large), their net earnings decrease with s_l except at $s_l = \underline{s_l}(f)$, $s_l = \overline{s_l}(f)$, or both depending on at which s_l the switch to a local job occurs, where the net earnings increase discontinuously from Lemma A4, if the above conditions are satisfied.

Net earnings of workers with $a < e^+(0)$ are maximized at $s_l = 0$, because their earnings when the return is negative decrease with s_l from Proposition 2 and thus the earnings when $s_l \to \overline{s_l}(f)$ from above are lower than the earnings at $s_l = 0$.

Net earnings of workers with $a \geq \min\{e^+(1), \overline{e}\}$ are maximized at either $s_l = 0$ or $s_l = \underline{s_l}(f)$, since the earnings increase discontinuously at $s_l = \underline{s_l}(f)$. From the proof of Lemma A4, net earnings of such workers with educational spending e at $s_l = 0$ equal

$$\left[\alpha A \left(\frac{F(e^+(0))\underline{h}_l}{\frac{\delta}{p}\left[\int_{e^+(0)}^{\overline{e}} ef(e)de + [1-F(\overline{e})]\overline{e}\right]}\right)^{1-\alpha} \frac{\delta}{p} - 1\right] e = \left[\alpha A \left(\frac{p}{\delta}\frac{1-\alpha}{\alpha}\frac{\underline{h}_l}{e^+(0)}\right)^{1-\alpha} \frac{\delta}{p} - 1\right] e, \tag{63}$$

where the equality sign is from (10) and (5).

From (14) in the proof of Proposition 1 and (2), their net earnings at $s_l = \underline{s_l}(f)$ equal

$$\left\{\alpha A \left[\frac{(1-\alpha)\left(\underline{h}_l + \frac{s_l e^+}{p}\right)}{\alpha\delta\frac{(1-s_l)e^+}{p}}\right]^{1-\alpha} \delta\frac{1-s_l}{p} - 1\right\} e = \left\{\alpha A \left[(1-\alpha)A\frac{s_l}{p}\right]^{\frac{1-\alpha}{\alpha}} \delta\frac{1-s_l}{p} - 1\right\} e \tag{64}$$

$$\leq \left\{\alpha A \left[(1-\alpha)A\frac{1-\alpha}{p}\right]^{\frac{1-\alpha}{\alpha}} \delta\frac{\alpha}{p} - 1\right\} e. \tag{65}$$

where the equality sign is from the fact that the return for local jobs is 0 at $s_l = \underline{s_l}(f)$.

From (63) and (65), the net earnings are maximized at $s_l = 0$ if

$$\left(\frac{p}{\delta}\frac{1-\alpha}{\alpha}\frac{h_l}{e^+(0)}\right)^{1-\alpha} > \left[(1-\alpha)A\frac{1-\alpha}{p}\right]^{\frac{1-\alpha}{\alpha}}\alpha, \tag{66}$$

which holds when A and δ are small and p is large, since $e^+(0)$ does not depend on A, δ, and p.

As for those who choose a national job when s_l is small and a local job when s_l is large, the proof for those who choose a local (national) job for any s_l applies when the switch to a local job occurs at $s_l \le \underline{s_l}(f)$ ($s_l \ge \overline{s_l}(f)$). When the switch occurs at $s_l \in (\underline{s_l}(f), \overline{s_l}(f))$, the proof of those who choose a national job for any s_l applies, because earnings when the return is negative decrease with s_l from Proposition 2.

(ii) From the above results, if $e < \min\{e^+(0), \alpha(p\underline{h}_l + E(e|e<e^+(0)))\}$ when $E(e|e<e^+(0)) \le \frac{1+\alpha}{1-\alpha}p\underline{h}_l$ or if $e < \min\{e^+(0), \Omega(e^+(0))\}$ when $E(e|e<e^+(0)) > \frac{1+\alpha}{1-\alpha}p\underline{h}_l$, workers choose a local job for any s_l, and their net earnings decrease with s_l except at $s_l = \overline{s_l}(f)$, where the earnings increase discontinuously. Net earnings of such workers are maximized at $s_l = 0$, because the earnings when the return is negative decrease with s_l from Proposition 2.

(iii) As for workers who have abundant wealth and thus choose a national job for any s_l, i.e. $a \ge \min\{e^+(1), \overline{e}\}$, their net earnings decrease with s_l for $s_l < \underline{s_l}(f)$ and $s_l > \overline{s_l}(f)$ from Proposition 2, increase (decrease) discontinuously at $s_l = \underline{s_l}(f)$ ($s_l = \overline{s_l}(f)$) from Lemma A4, and increase (decrease) with s_l for $s_l \in (\underline{s_l}(f), s_l^\flat]$ ($s_l \in [s_l^\sharp, \overline{s_l}(f))$), if $e^+(0) > \frac{\alpha p\underline{h}_l}{1-\alpha}$, $E(e|e<e^+(0)) > \alpha \max\left\{\frac{p\underline{h}_l}{1-\alpha}, \frac{e^+(0)}{1+\alpha}\right\}$, and $s_l^\flat > \underline{s_l}(f)$, where $s_l^\flat \in (0, s_l^\sharp)$ satisfies $s_l^\flat = (1-\alpha) - \alpha\frac{p\underline{h}_l}{E(e|e<e^+(s_l^\flat))}$, from Proposition A1 (i). ($s_l^\flat, s_l^\sharp < 1-\alpha < \overline{s_l}(f)$ from Proposition A1 (i) and Lemma A1 in the proof of Proposition 1.) The condition holds when A and δ are large or p is small, because $e^+(0)$ does not depend on these parameters, $s_l^\flat$ decreases with p (since $e^+(s_l)$ decreases with p, as shown above), and $\underline{s_l}(f)$ increases with p and decreases with A and δ from Proposition 1.

Their net earnings are maximized at either $s_l = 0$ or s_l satisfying $\frac{d(w_n h_n)}{ds_l} = 0$, where the latter s_l satisfies $s_l \in (s_l^\flat, 1-\alpha)$ and thus $s_l < 1-\alpha < \overline{s_l}(f)$ from Proposition A1 (i) and Lemma A1 in the proof of Proposition 1.

From (63) and (64) above, net earnings of such workers with educational spending e at $s_l = 0$ is smaller than their net earnings at $s_l = s_l^\flat$ (and thus the earnings at s_l satisfying $\frac{d(w_n h_n)}{ds_l} = 0$) if $\left(\frac{\underline{h}_l}{e^+(0)}\right)^{1-\alpha} < \left[\frac{1}{e^+(s_l^\flat)}\left(\underline{h}_l + \frac{s_l^\flat e^+(s_l^\flat)}{p}\right)\right]^{1-\alpha}(1-s_l^\flat)^\alpha$, which holds if

$$\left(\frac{\underline{h}_l}{e^+(0)}\right)^{1-\alpha} < \left[\frac{1}{\overline{e}}\left(\underline{h}_l + \frac{s_l^\flat \overline{e}}{p}\right)\right]^{1-\alpha}(1-s_l^\flat)^\alpha. \tag{67}$$

The condition holds when p is small enough, because $e^+(0)$ does not depend on A, δ, and p, $s_l^\flat$ decreases with p, and the RHS increases with $s_l^\flat$:

$$\frac{1-\alpha}{\underline{h}_l + \frac{s_l^\flat \overline{e}}{p}}\frac{\overline{e}}{p} - \frac{\alpha}{1-s_l^\flat} = \frac{(1-\alpha-s_l^\flat)\overline{e} - \alpha p\underline{h}_l}{\left(\underline{h}_l + \frac{s_l^\flat \overline{e}}{p}\right)p(1-s_l^\flat)} = \alpha p\underline{h}_l\frac{\frac{e^+(s_l^\flat)}{E(e|e<e^+(s_l^\flat))} - 1}{\left(\underline{h}_l + \frac{s_l^\flat \overline{e}}{p}\right)p(1-s_l^\flat)} > 0. \tag{68}$$

As for workers who have limited wealth and thus choose a local job for any s_l, i.e. $a < e^+(0)$, their net earnings decrease with s_l for $s_l < \underline{s_l}(f)$ and $s_l > \overline{s_l}(f)$ from Proposition 2, decrease (increase) discontinuously at $s_l = \underline{s_l}(f)$ ($s_l = \overline{s_l}(f)$) from Lemma A4, and increase (decrease) with s_l for $s_l \in (\underline{s_l}(f), \min(s^{\triangle}_{l,h}(e), \overline{s_l}(f))]$ ($s_l \in [s^{\triangledown}_{l,h}(e), \overline{s_l}(f))$ when $s^{\triangledown}_{l,h}(e) < \overline{s_l}(f)$), if $s^{\triangle}_{l,h}(e) > \underline{s_l}(f)$ and $e > \max\{\alpha(p\underline{h_l}+e^+(0)), \Lambda(\overline{e})\}$ from Proposition A1 (ii). The condition holds when A and δ are large, because $e^+(0)$, $\Lambda(\overline{e})$, and $s^{\triangle}_{l,h}(e)$ do not depend on these parameters, and $\underline{s_l}(f)$ decreases with A and δ. The condition holds when p is small, because $\Lambda(\overline{e})$ increases with p, $s^{\triangle}_{l,h}(e) > [s_l$ satisfying $s_l = \frac{(1-\alpha)-(1+\alpha)\frac{ph_l}{e^+(s_l)}}{2}] > 0$ when p is small enough, while $\underline{s_l}(f)$ approaches 0 as p decreases, which is from $\underline{s_l}(f)$ being decreasing in p, and from (14) and Lemma A1 in the proof of Proposition 1.

Their net earnings are maximized at either $s_l = 0$, or $s_l \in (s^{\triangle}_{l,h}(e), \overline{s_l}(f))$ such that $\frac{d(w_l h_l)}{ds_l} = 0$. (The earnings at $s_l = \overline{s_l}(f)$ are smaller than the earnings when $s_l \to \overline{s_l}(f)$ from above and thus cannot be maximum.) From the proof of Lemma A4, net earnings of such workers at $s_l = 0$ equal

$$(1-\alpha)A\left(\frac{\frac{\delta}{p}\left[\int_{e^+(0)}^{\overline{e}} ef(e)de + [1-F(\overline{e})]\overline{e}\right]}{F(e^+(0))\underline{h_l}}\right)^{\alpha}\underline{h_l} = (1-\alpha)A\left(\frac{\delta}{p}\frac{\alpha}{1-\alpha}\frac{e^+(0)}{\underline{h_l}}\right)^{\alpha}\underline{h_l}. \tag{69}$$

From (14) in the proof of Proposition 1, net earnings of such workers with e at $s_l = s^{\triangle}_{l,h}(e)$ equal $(1-\alpha)A\left[\frac{\alpha\delta\frac{(1-s^{\triangle}_{l,h}(e))e^+(s^{\triangle}_{l,h}(e))}{p}}{(1-\alpha)\left(\underline{h_l}+\frac{s^{\triangle}_{l,h}(e)e^+(s^{\triangle}_{l,h}(e))}{p}\right)}\right]^{\alpha}\left(\underline{h_l}+\frac{s^{\triangle}_{l,h}(e)}{p}e\right) - e$. The net earnings at $s_l = s^{\triangle}_{l,h}(e)$ is greater than at $s_l = 0$ if

$$(1-\alpha)A\left(\frac{\delta}{p}\frac{\alpha}{1-\alpha}\right)^{\alpha}\left\{\left[\frac{(1-s^{\triangle}_{l,h}(e))e^+(s^{\triangle}_{l,h}(e))}{\underline{h_l}+\frac{s^{\triangle}_{l,h}(e)e^+(s^{\triangle}_{l,h}(e))}{p}}\right]^{\alpha}\left(\underline{h_l}+\frac{s^{\triangle}_{l,h}(e)}{p}e\right) - \left(\frac{e^+(0)}{\underline{h_l}}\right)^{\alpha}\underline{h_l}\right\} - e > 0, \tag{70}$$

which is true when A or δ is large enough, because the expression inside the curly bracket is positive from $\frac{d(w_l h_l)}{ds_l} > 0$ for $s_l \le s^{\triangle}_{l,h}(e)$. It is true when p is small enough too, since the expression is positive for $s_l \le s^{\triangle}_{l,h}(e)$, and, as shown above, $s^{\triangle}_{l,h}(e)$ does not converge to 0 as p decreases.

Finally, as for workers who choose a national (local) job when s_l is small (large), i.e. $a \in [e^+(0), \min\{e^+(1), \overline{e}\})$, the result is clearly similar to workers who choose a national (local) job for any s_l, when the shift to a local job occurs at $s_l > \overline{s_l}(f)$ ($s_l < \underline{s_l}(f)$). When the shift occurs at $s_l \in [\underline{s_l}(f), \overline{s_l}(f)]$, their net earnings increase discontinuously at $s_l = \underline{s_l}(f)$ and $s_l = \overline{s_l}(f)$, and they are maximized at either $s_l = 0$ or $s_l \in (\underline{s_l}(f), \overline{s_l}(f))$ satisfying $\frac{d(w_n h_n)}{ds_l} = 0$, $\frac{d(w_l h_l)}{ds_l} = 0$, or $a = e^+(s_l)$ (s_l at which the switch occurs). (The earnings at $s_l = \overline{s_l}(f)$ are smaller than the earnings when $s_l \to \overline{s_l}(f)$ from above and thus cannot be maximum.) From the argument above, the net earnings are maximized at the latter s_l when A and δ are large enough or p is small enough.

■