Is bilingual education desirable in multilingual countries?

Kazuhiro Yuki*

August 2022

Abstract

In many countries, people use their mother tongue in local business, but use the language of the former colonizer in national business. How much weight should be placed on teaching one's mother tongue and the lingua franca is a critical issue in these countries.

This paper develops a model to examine these issues theoretically. It is shown that balanced education of the two languages is critical for skill development of those with limited wealth. It is also found that balanced bilingual education yields higher earnings net of educational expenditure than lingua-franca-only education *only when* a country has favorable educational and technological conditions (productivity is reasonably high and education is reasonably effective) and *only for* those with adequate wealth. Policy implications of the results are also discussed.

Keywords: language policy, bilingual education, human capital, economic development
JEL classification numbers: I25, J24, O15, Z13

*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. Invaluable comments from an anonymous referee and 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. Data sharing is not applicable to this paper as no new data were created or analyzed in this study.

1 Introduction

Many countries in sub-Saharan Africa and countries such as India and the Philippines are populated by multiple ethnic groups who use their mother tongue in daily life and in local business, but use the language of the former colonizer as the lingua franca (common language) in national business and in communications with other groups. In these countries, how much weight should be placed on teaching one's mother tongue and on teaching the lingua franca and which language should be used as a language of instruction of other subjects are critical issues.[1,2] Acquiring the skill to use one's mother tongue is less demanding because a part of the skill is taught at home, but its use is limited to the ethnic community. By contrast, acquiring the skill to use the lingua franca is harder, but the skill is important in many modern sector jobs.

Students have little choice between mother tongue and lingua franca education in basic education, i.e., primary and lower secondary education, because relative weights of the two types of education are mostly determined by the government. In sub-Saharan Africa, former French colonies had maintained French as the sole language of instruction and former British colonies had conducted mother tongue education partially, although Francophone countries began using ethnic languages and many Anglophone countries reduced the weight on mother tongue education recently (Albaugh, 2007; Heugh, 2011a).

A general consensus among specialists on language and education is that placing emphasis on mother tongue education 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 toward lingua franca education (Heugh, 2011a).

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 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 mother tongue education for skill development by experts, many parents in sub-Saharan Africa are resistant to mother tongue education because they believe that it does not help their children get a job (Albaugh, 2007). Indeed, the economic return to a lingua franca in multilingual countries can be large: Azam, Chin, and Prakash (2013) find that the return to speak fluent English, a lingua franca in India, is as large as the return to secondary education and half as large as the return to undergraduate education.

The goal of this paper is to develop a model and examine the above-stated issues theoretically.

Model: In the model, two kinds of "jobs", called *national jobs* and *local jobs*, requiring different types of skill exist and the final good is produced from them. In the real economy, national jobs

[1]The same issues are also relevant to many small nations, including ethnolinguistically homogenous nations, in which the lingua franca in business is English because of strong dependence on international business.

[2]A similar conflict arises in basic education of low-income countries in general, including monolingual ones, between teaching vocational skills that are directly useful in local jobs (e.g., farming and related skills in an agrarian community) and teaching academic skills that are important in jobs involving modern business practice and technology. Results of the paper apply to the issue of relative weights on the education of these skills as well.

correspond mainly to jobs in the modern sector (the government and a part of the private sector using modern technology), while local jobs correspond mainly to jobs in the traditional sector (traditional agriculture, the urban informal sector, and the household sector).

Education is costly and allows the individuals to acquire the job-specific skill for each type of job. The skill key to national jobs is the knowledge of the lingua franca, while the use of the mother tongue in education is the skill key for local jobs.

The individual chooses the amount of educational spending, but *cannot* choose its allocation over the development of the two types of skill, which is fixed reflecting the fact that relative weights of the two types of education are mostly determined by the government. The level of skill for local jobs is *positive* without education (i.e., a portion of the mother tongue skill is taught by family members), while the level of skill for national jobs is zero without education.

The paper mainly focuses on the case in which some individuals do not have enough wealth to make optimal educational investment, reflecting the fact that, in many developing countries, students must rely on family wealth to pay for study materials, commuting cost, and others even when public schools do not charge tuitions. Because the level of skill 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: The paper examines how a change in relative weights of the two types of education affects educational and job choices and earnings. Main results can be summarized as follows.

First, balanced education of the two types of skill is critical for skill development of those who have limited wealth and thus choose a local job: when the allocation of educational spending is very biased, the return to education becomes negative for them and they do not spend on education.[3]

Second, balanced bilingual education yields higher earnings net of educational expenditure than lingua-franca-only education *only when* the country has good educational and technological conditions (i.e., productivity is reasonably high and education is reasonably effective in skill development) and *only for* those with sufficient wealth. Net earnings of those with little wealth *decrease* with the weight on mother tongue education and are highest under lingua-franca-only education. This is true for *everyone* under unfavorable conditions.[4] In the real economy, the educational and technological conditions tend to be related to the country's level of economic and social development. Hence, the result suggests that if the level of development is reasonably high, the balanced education is economically desirable except for the very poor; otherwise, lingua-franca-only education is desirable in terms of the economic outcome. By contrast, education biased toward mother tongue education leads to low net earnings.[5]

[3]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).

[4]Further, numerical simulations suggest that, when the proportion of those with limited wealth is very high, lingua-franca-only education maximizes net earnings of all.

[5]It is also found that bilingual education with a *very small* weight on mother tongue education is *worse* than

The results imply that a *trade-off* between educational and economic outcomes always exists for the very poor and under unfavorable educational and technological conditions, for all, when the former is measured by the mother tongue skill, which is an essential skill in daily life even for those with a national job: under lingua-franca-only (balanced bilingual) education, their net earnings are highest (low) but their mother tongue skill is lowest (high).

Policy implications: The results have the following policy implications. When the educational and technological conditions are favorable, the government that takes into account both educational and economic outcomes of individuals would implement balanced bilingual education *together with* a redistributive policy that enables those with little wealth to expend sufficiently more on education so that they benefit economically from the balanced education. By contrast, when the educational and technological conditions are not good, the government that balances the educational outcome against the economic outcome would choose bilingual education with a *smaller* (*but not too small*) weight on mother tongue education than under the more favorable conditions (together with redistribution toward the very poor).

Note that the model does not consider possibly important effects the choice of languages in education has on social capital, political participation, national unity, and public goods provision. Policy implementation in the actual society needs to take into account these effects as well.

Related literature: To the author's knowledge, this paper, along with Yuki (2021) in which the common language is the mother tongue of the dominant group, is the first attempt to examine theoretically how relative weights of the two types of education affect educational investment and net earnings of individuals with different wealth. There exist works examining the issue empirically and works analyzing related issues theoretically.

The effect of language policy in education on academic achievement of students is studied extensively in education and lingustics (Heugh, 2011a; Baker and Wright, 2017) and slightly in economics (Angrist, Chin and Goody, 2008; Ramachandran, 2017). The empirical findings are consistent with the model's result on the educational outcome.

Very few studies examine labor market outcomes. Angrist and Lavy (1997) find that replacing French with Arabic as the medium of instruction in post-primary education greatly lowered returns to schooling in Morocco. Consistent with the model's result on earnings, the finding suggests that a significant increase in the weight on mother tongue education lowered wages in the developing country. Chakraborty and Bakshi (2016) show that the policy change in the Indian state of West Bengal that abolished English education in primary schools has a negative effect on wages. This is consistent with the model's result that education biased toward the mother tongue skill results in low earnings. Cappellari and Di Paolo (2018) find a positive effect on earnings of a bilingual education reform in the Catalonia region of Spain that greatly increased the weight on Catalan in mandatory education. The finding is not relevant for the present paper because the mother tongue of many students (migrants from Spanish-speaking regions and their offspring) was Spanish.

lingua-franca-only education: a switch from the latter to the former does not improve the educational outcome of students from poor families and lowers net earnings of all.

Pool (1991) examines the choice of official language(s) in a multilingual society in which earnings are exogenous, adopting an official language is costly for the nonnative, and translation among multiple official tongues is costly and financed by tax. He shows that there exists an efficient and fair choice of official language(s), if proper inter-group redistribution is conducted. Ortega and Tangerås (2008) develop a model of a society of two language groups without intra-group heterogeneity, in which the dominant group determine the type(s) of schools (monolingual in either language or bilingual) accessible to each group, individuals decide whether to attend school, and goods are produced from bilateral random matching only when pairs speak the same language. They show that the dominant group choose laissez-faire or restrict access to schools using the language of the subordinate group, while the subordinate prefer schools using their mother tongue.[6]

Besides examining a different issue, the present work differs from these works in the following respects. First, in this work, the common language is not a mother tongue of any group and individuals are heterogenous in wealth available for education, while in the preceding works, official or education language(s) are native languages of either group(s) and individuals are homogenous within each group. This paper adopts different settings, because it focuses on developing countries where the common language is typically the language of the former colonizer and family wealth is a critical determinant of educational investment, whereas the existing works mainly focus on developed countries. Second, unlike the other works, the present paper does not take into account the effect of the size of language groups, such as network externalities in language usage. Finally, unlike Ortega and Tangerås (2008), educational institutions are given rather than determined endogenously and strategic interactions among agents do not exist.

Organization of the paper: Section 2 presents the model, and Section 3 examines educational and job choices of individuals. Section 4 provides a preliminary analysis of how a change in relative weights of the two types of education affects individual choices and earnings, and Section 5 presents the main results and discusses their policy implications. Section 6 concludes. Appendices A and B present auxiliary results, and Appendix C contains proofs of lemmas and propositions.

2 Model

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

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

[6] Other works studying issues related to language in economics include Lazear (1999), Ginsburgh, Ortuño-Ortín, and Weber (2005), Clots-Figueras and Masella (2013), Desmet, Ortuño-Ortín, and Wacziarg (2012), and Galor, Özak, and Sarid (2018). In political science, works such as Laitin (1992) and Albaugh (2007) examine political aspects of the choice of education languages in multilingual societies.

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

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

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

Each person has a wealth endowment to expend on education for developing the skills required in the two types of jobs. The skill key to national jobs is the knowledge of the lingua franca, while the use of the mother tongue in education is the skill key for local jobs.[7] The model would be relevant to many countries in sub-Saharan Africa and countries such as India and the Philippines in which the lingua franca in national business is the language of the former colonizer.[8]

She chooses the amount of educational spending e, which largely depends on years of schooling in the real world, but *cannot* choose its allocation over the development of the two types of skill, which is fixed reflecting the fact that relative weights of lingua franca and mother tongue education are mostly determined by the government in basic education, i.e., primary and lower secondary education.

The human capital production functions of the two types of skill are:[9]

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

$$h_n(e, s_l) = \delta_n (1 - s_l) e, \tag{4}$$

$$\text{where } s_l \in [0,1],\ \underline{h}_l, \delta_l, \delta_n > 0, \text{ and } e \leq \bar{e}.$$

In the above equations, $\underline{h}_l$ is the level of skill for local jobs when $e = 0$, $s_l \in [0,1]$ is the proportion of e allocated to the development of the skill for local jobs, and δ_l (δ_n) is the productivity of the education technology of the skill for local (national) jobs.[10]

[7] An alternative interpretation, which would apply to low-income countries in general, including monolingual ones, is that the skill for local jobs corresponds to vocational skills that are directly useful in local jobs (e.g., farming and related skills in an agrarian community) and the skill for national jobs corresponds to academic skills that are important in jobs involving modern business practice and technology.

[8] For analytical tractability, the model assumes that either ethnic groups are symmetric in every respect or workers of different groups with a given type of jobs are perfectly substitutable in production. The former assumption would be relevant to countries in which a single dominant ethnic group does not exist. The model would also be relevant to many small nations, including ethnolinguistically *homogenous* nations, in which the lingua franca in business is English because of strong dependence on international business.

[9] For analytical tractability, the human capital production functions do not assume complementarities between the two types of education, in particular, a positive effect of mother tongue education on the development of the lingua franca skill, which is empirically plausible (Heugh, 2011a). As explained in footnote 28 of Section 5.1, assuming the complementarity would hardly affect the main results.

[10] $\delta_n < \delta_l$ would be reasonable considering a higher cost effectiveness of mother tongue education in skill development (Heugh, 2011b).

The level of skill for local jobs, i.e., the mother tongue skill, is positive without education reflecting the fact that the skill can be developed partly at home, while the level of skill for national jobs, i.e., the lingua franca skill, is zero without education. The production functions are assumed to be linear to make the model analytically tractable.[11] Because the marginal return to educational investment does not depend on e, i.e., $w_n\delta_n(1-s_l)-1$ for national jobs and $w_l\delta_l s_l - 1$ for local jobs, the upper limit $\overline{e}$ is set so that realized e does not become too large for some individuals.[12]

Although the case in which no one faces the wealth constraint on educational investment is also analyzed, the default setting is that some individuals do not have enough wealth to make optimal investment. This reflects the fact that, in many developing countries, students must rely on family wealth to pay for study materials, commuting cost, uniforms, and supplementary education even when public schools do not charge tuitions. A person who has wealth (endowment) a 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, together with wealth net of educational spending $a-e$, are spent on final good consumption.

3 Educational and job choices

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

3.1 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 $a \leq \overline{e}$ 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 $w_n h_n(e, s_l) < w_l h_l(e, s_l)$ holds for $e < e^+$, i.e., those who spend $e = a < e^+$ on education choose a local job. When $e^+ < \overline{e}$, those who spend $e > e^+$ choose a national job, while when $e^+ = \overline{e}$, those who spend $e = \overline{e}$ are indifferent between the two types of jobs. Figure 1 illustrates how educational and job choices and earnings depend on wealth a when $e^+ < \overline{e}$.[13] Intuitively speaking, those who have limited wealth and thus cannot spend much on education choose a local job, because a part of the skill for the job ($\underline{h}_l$) does not require educational spending.

[11] When human capital production functions exhibit diminishing marginal returns to educational spending, the maximum level of e for each type of jobs is determined endogenously, which complicates equations determining equilibrium significantly and, compared to the linear specification, increases the number of qualitatively distinct cases to analyze. Further, the relationship between the distribution of wealth and realized cases becomes much more complicated than the present setting, under which the relationship is illustrated by Figure 3 below.

[12] Note that the return to educational investment for national jobs is *always strictly* positive, because $w_n h_n(e, s_l) - e = [w_n h_n(1, s_l) - 1]\, e \geq w_l h_l(0, s_l) = w_l \underline{h}_l > 0$ must hold for individuals choosing a national job. This implies that, without the upper limit $\overline{e}$, even a very wealthy individual spends her entire wealth on education.

[13] 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.

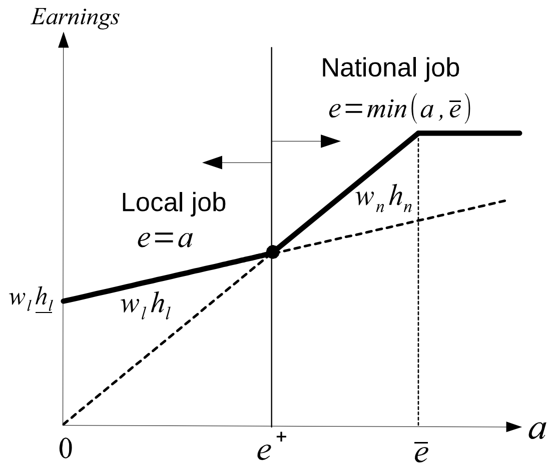


Figure 1: Educational and job choices and earnings when education is worthwhile for all jobs and $e^+ < \overline{e}$

When $e^+ < \overline{e}$, the aggregate human capital of workers in local and national jobs, H_l and H_n, are given by

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

where e^+ is determined by

$$w_n h_n(e^+, s_l) = w_l h_l(e^+, s_l) \quad (6)$$

$$\Leftrightarrow \alpha H_l(e^+, s_l) h_n(e^+, s_l) = (1 - \alpha) H_n(e^+, s_l) h_l(e^+, s_l) \ \text{ (from (2))}. \quad (7)$$

When $e^+ = \overline{e}$, H_l and H_n are given by

$$H_l(\pi_n, s_l) = \int_0^{\overline{e}} h_l(e, s_l) f(e) de + (1 - \pi_n) [1 - F(\overline{e})] h_l(\overline{e}, s_l) \quad (8)$$

$$\text{and } H_n(\pi_n, s_l) = \pi_n [1 - F(\overline{e})] h_n(\overline{e}, s_l), \quad (9)$$

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

$$\pi_n = \alpha \left\{ 1 + \frac{\int_0^{\overline{e}} h_l(e, s_l) f(e) de}{[1 - F(\overline{e})] h_l(\overline{e}, s_l)} \right\}. \quad (10)$$

3.2 When education is not worthwhile for local jobs

Next, consider the case in which education is not worthwhile, i.e., the return to educational investment is negative, for local jobs. (Education must always be worthwhile for national jobs, because the skill 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 individuals with $a < e^+$ do not spend on education and choose a local job. When $e^+ < \overline{e}$, those with $a > e^+$ choose a national job, whereas when $e^+ = \overline{e}$,

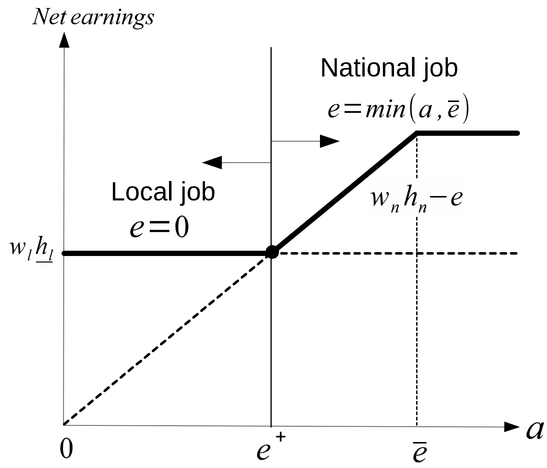


Figure 2: Educational and job choices and earnings net of educational spending when education is not worthwhile for local jobs and $e^+ < \overline{e}$

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 a when $e^+ < \overline{e}$.

In this case, H_l and H_n when $e^+ < \overline{e}$ are given by

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

where e^+ is determined by the following equation with $H_l = H_l(e^+, s_l)$ and $H_n = H_n(e^+, s_l)$:

$$w_n h_n(e^+, s_l) - e^+ = w_l h_l(0, s_l) \Leftrightarrow [w_n h_n(1, s_l) - 1] e^+ = w_l \underline{h}_l$$
$$\Leftrightarrow \left[\alpha A \left(\frac{H_l}{H_n}\right)^{1-\alpha} \delta_n (1 - s_l) - 1\right] e^+ = (1-\alpha) A \left(\frac{H_n}{H_l}\right)^{\alpha} \underline{h}_l \text{ (from (2))}. \tag{12}$$

When $e^+ = \overline{e}$, H_l and H_n are given by

$$H_l(\pi_n, s_l) = \{F(\overline{e}) + (1 - \pi_n)[1 - F(\overline{e})]\} \underline{h}_l \text{ and } H_n(\pi_n, s_l) = \pi_n [1 - F(\overline{e})] h_n(\overline{e}, s_l), \tag{13}$$

where π_n is the solution to (12) with $H_l = H_l(\pi_n, s_l)$ and $H_n = H_n(\pi_n, s_l)$.

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

So far, the individual choices and the determination of several 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 (see Appendix B for a fully detailed statement).

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 individuals 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*), δ_n *and* δ_l (*respectively, effectiveness of education of the skill for national jobs and for local jobs*).

(*ii*) $e^+ < \bar{e}$ *holds if* $F(\bar{e})$ *is large, and* $e^+ = \bar{e}$ *holds otherwise. When* $F(\bar{e}) \leq 1 - \alpha$, $e^+ = \bar{e}$ *always holds.*

The first part of the proposition shows that, when the proportion of educational spending allocated to the development of the skill 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 $w_l \delta_l s_l - 1$ is negative because the marginal effect of e on the human capital for local jobs, $\delta_l s_l$, is very small, which dominates high w_l due to large $\frac{H_n}{H_l}$ (total human capital in national jobs relative to the one in local jobs). When s_l is very high, the return is negative because w_l is very low due to small $\frac{H_n}{H_l}$, which dominates a large marginal effect of e on the human capital.[14] 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., $e = a$. Increases in TFP and in effectiveness of education in skill development widen the range of s_l over which education is worthwhile for those choosing a local job, because the wage rate or the marginal effect of educational expenditure on the human capital increase.

The result that those who choose 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 poorly, either because what they learn is mostly irrelevant to future jobs in the local or ethnic community (when s_l is very low) or because their future earnings are low due to deficient skill of workers in complementary modern sector jobs (when s_l is very high).

[14] Assuming a positive effect of mother tongue education [education of the skill for local jobs] on the development of the lingua franca skill [the skill for national jobs] (footnote 9) would not affect the result qualitatively. When s_l is very small, this is because the negative return to education for local jobs of the proposition is due to a small marginal effect of e on the skill for *local jobs*. When s_l is very large, the negative return of the proposition is due to low w_l, which results from small $\frac{H_n}{H_l}$. Because very large s_l implies a very small weight on the education of the skill for national jobs, it is almost certain that the direct negative effect of large s_l on h_n outweighs the positive indirect effect on h_n through the education of the skill for local jobs. Hence, $\frac{H_n}{H_l}$ remains small even when the indirect effect is taken into account.

The result shows that balanced education of the skill for national jobs and the skill for local jobs (moderate s_l) is critical for skill development of those with limited wealth. This is in line with a general consensus among specialists on language and education that placing emphasis on mother tongue education at least in primary education is effective for students to acquire adequate skills (Heugh, 2011a). It is also consistent with empirical findings in economics (Angrist, Chin and Goody, 2008; Ramachandran, 2017).[15]

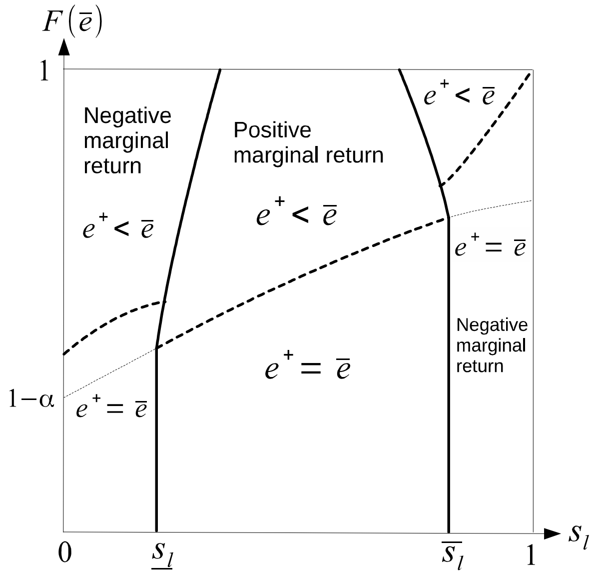


Figure 3: Proposition 1

The second part of the proposition states that $e^+ < \overline{e}$ holds, that is, some of those who cannot afford the optimal level of education $\overline{e}$ get a national job, if their share $F(\overline{e})$ is high; 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 when $\int_0^{\overline{e}} ef(e)de < (1-\alpha)\overline{e}$.[16] 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.) For given s_l, $e^+ < \overline{e}$ ($e^+ = \overline{e}$) holds when $F(\overline{e})$ is relatively high (low).[17] Note that $e^+ = \overline{e}$ holds for any s_l when $F(\overline{e}) \leq 1-\alpha$, which is used in the later analysis.

Figure 4 summarizes how educational and job choices depend on s_l and a when $F(\overline{e})$ is high

[15] 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 the skills. Ramachandran (2017) finds that the educational reform in Ethiopia which introduced mother tongue instruction in primary education has positive effects on the reading skill and education.

[16] 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 $F(\overline{e}) = 1-\alpha$. The main results are unchanged in these cases.

[17] 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 proved in Claim 1 of Appendix A.

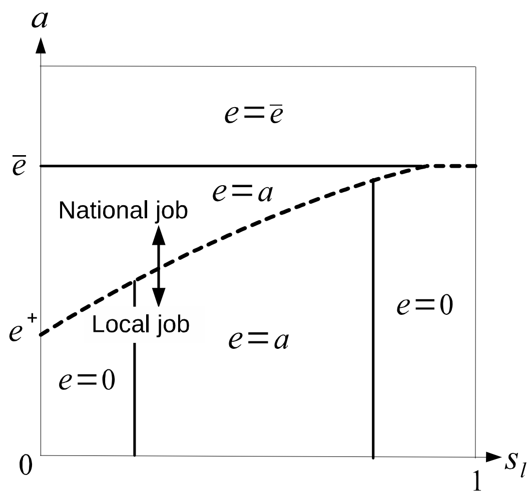


Figure 4: Dependence of educational and job choices on s_l and a

enough that $e^+ < \bar{e}$ holds for not high s_l (see Figure 3).[18] When s_l is very high or very low, individuals who have wealth $a < e^+$ and thus take a local job do not spend on education, while when s_l is at an intermediate level, they choose $e = a$. For any s_l, those who have $a > e^+$ and thus take a national job spend $e = \min\{a, \bar{e}\}$.

4 Preliminary Analysis of effects of s_l

This section provides a preliminary analysis of how a change in relative weights of the two types of education affects job choices 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 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. Based on the result in this section, the next section presents the main results by taking into account the proposition.

4.1 Effects on job choices and wage rates

The next lemma examines the effect of s_l on the variables governing job choices, i.e., 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* $e^+ = \bar{e}$, $\dfrac{d\pi_n}{ds_l} < 0$.

The higher the proportion of educational spending allocated to the development of the skill for local jobs, the higher the fraction of individuals choosing 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. For a given value of the variable governing a job choice, i.e., e^+ when $e^+ < \bar{e}$ and π_n when $e^+ = \bar{e}$, 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

[18]The line for e^+ is upward-sloping when $e^+ < \bar{e}$ because as shown in Lemma 1 below, e^+ increases with s_l.

job, while increased s_l raises w_n and lowers w_l through a negative effect on the aggregate human capital ratio $\frac{H_n}{H_l}$, which induces the shift of workers in the opposite direction. When $\underline{h}_l > 0$, the former effect dominates and thus a higher fraction of workers choose a local job.

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

4.2 Effects on earnings

Hence, an increase in s_l raises (lowers) the human capital of those who choose a local (national) job but lowers (raises) their wage rate. Which effect dominates? The following propositions examine the effect of s_l on earnings based on the lemmas. Proposition 2 examines the case in which the return to educational investment for local jobs is negative (the outer regions of Figure 3).

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

When s_l is small or large enough that the return to education for local jobs is negative, those who choose a local job due to limited wealth ($a < e^+$) do not spend on education (Proposition 1 (i)). The proposition shows that, under such situation, earnings of *everyone* decrease when a greater proportion of educational expenditure is allocated to the development of the skill for local jobs, i.e., the mother tongue skill. Earnings of workers with a local job decrease because their wage rate w_l falls due to lowered $\frac{H_n}{H_l}$ [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 workers with a national job fall because higher s_l lowers their human capital, which dominates a positive effect on their wage rate w_n.[19]

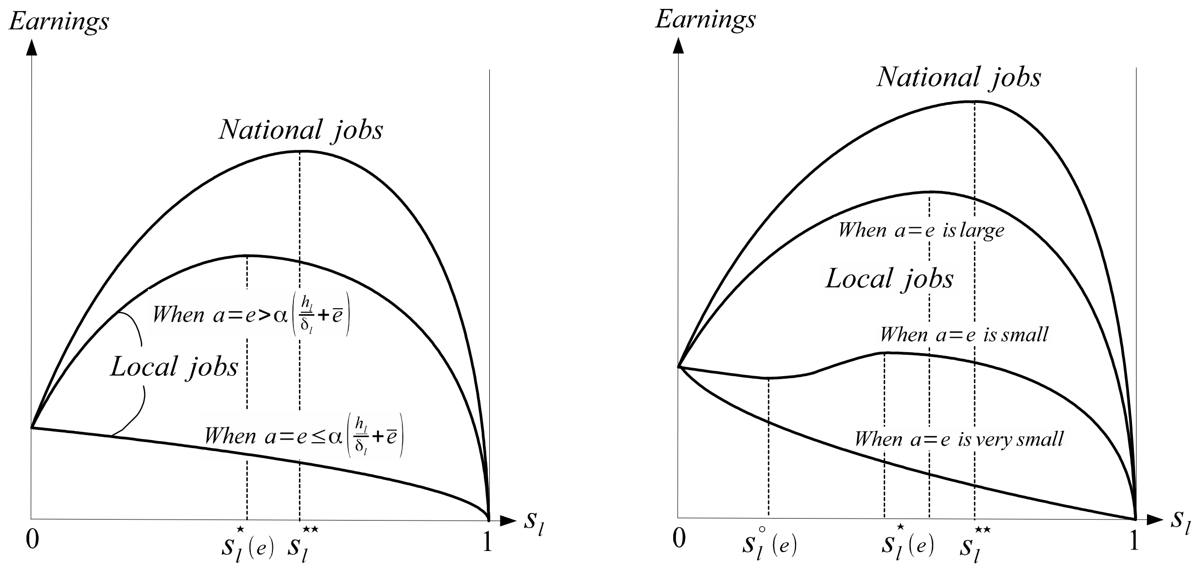
Proposition 3 examines the case in which the return is positive and $e^+ = \bar{e}$ holds (the lower-middle region of Figure 3), assuming, for ease of presentation, that this case exists for any s_l.

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

(i) Net earnings of workers with a national job increase (decrease) with s_l for $s_l < (>) s_l^{\star\star} \equiv (1-\alpha) - \alpha\frac{h_l}{\delta_l \bar{e}}$ and are highest at $s_l = s_l^{\star\star}$.[20]

[19] Assuming a positive effect of the education of the skill for local jobs on the development of the skill for national jobs (footnote 9) would not largely affect the qualitative result. In the model with such effect, an increase in s_l raises $\frac{H_l}{H_n}$ and thus w_n less than the original model. Hence, earnings of workers with a national job could increase with s_l only when the effect of s_l on the skill for national jobs h_n is positive. This is the case if the direct negative effect of s_l on h_n is dominated by the positive indirect effect on the skill through the education of the skill for local jobs. However, this is very unlikely when s_l is very large (footnote 14). By contrast, when s_l is very small, if the indirect effect is large and outweighs the direct effect, their earnings might increase with s_l.

[20] $s_l^{\star\star} \equiv (1-\alpha) - \alpha\frac{h_l}{\delta_l \bar{e}} > 0 \Leftrightarrow \bar{e} > \frac{\alpha}{1-\alpha}\frac{h_l}{\delta_l}$ is assumed thereafter.



(a) When $\overline{e} \leq \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$ (b) When $\overline{e} > \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$

Figure 5: Effect of s_l on earnings when the return is positive and $e^+ = \overline{e}$ (Proposition 3)

(ii) (a) If $\overline{e} \leq \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$*, net earnings of workers with a local job and wealth* $a = e > \alpha\left(\frac{h_l}{\delta_l} + \overline{e}\right)$ *increase (decrease) with* s_l *for* $s_l < (>) s_l^\star(e)$ *and are highest at* $s_l = s_l^\star(e)$*, where* $s_l^\star(e) < s_l^{\star\star}$ *and* $s_l^{\star\prime}(e) > 0$*,*[21] *while net earnings of those with smaller wealth decrease with* s_l*.*

(b) Otherwise, net earnings of workers with a local job and an intermediate level of $a (= e)$ *decrease with* s_l *for* $s_l < s_l^\circ(e)$*, increase with* s_l *for* $s_l \in (s_l^\circ(e), s_l^\star(e))$*, and decrease with* s_l *for* $s_l > s_l^\star(e)$*, where* $s_l^{\circ\prime}(e) < 0$*.*[22] *For those with large or very small* $a (= e)$*, the result is same as (a).*

Figure 5 (a) and (b) illustrate the proposition when $\overline{e} \leq \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$ and when $\overline{e} > \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$ respectively. As mentioned above, the proposition summarizes the effect of s_l on earnings, *assuming*, for ease of presentation, that the return is positive and $e^+ = \overline{e}$ holds for any s_l, although this is *not true*, as shown in Proposition 1: when s_l is very low or very large, the return is negative. When the main results are presented in the next section, Proposition 1 is taken into account.

In both cases, earnings of workers with a national job (those with $a \geq e^+ = \overline{e}$) increase with s_l for $s_l < s_l^{\star\star}$, decrease with s_l for $s_l > s_l^{\star\star}$, and are highest at $s_l = s_l^{\star\star}$. Intuitively, a positive expenditure on the education of the skill 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. An increase in s_l lowers their human capital $h_n(\overline{e}, s_l)$ but raises the wage rate w_n. When

[21] $s_l^\star(e)$ ($s_l^\circ(e)$ of (b)) is the greater (smaller) solution of $-e\overline{e}(s_l)^2 + \left[(1-\alpha)\overline{e} - (1+\alpha)\frac{h_l}{\delta_l}\right] e s_l + \left[-\alpha\left(\frac{h_l}{\delta_l} + \overline{e}\right) + e\right]\frac{h_l}{\delta_l} = 0$. $s_l^\star(\overline{e}) = s_l^{\star\star}$ holds.

[22] To be precise, when $a = e \in \left(\frac{\alpha\left(\frac{h_l}{\delta_l} + \overline{e}\right)}{1 + \frac{\delta_l}{4h_l\overline{e}}\left[(1-\alpha)\overline{e} - (1+\alpha)\frac{h_l}{\delta_l}\right]^2}, \alpha\left(\frac{h_l}{\delta_l} + \overline{e}\right)\right)$.

s_l is low (high), the latter effect dominates (is dominated by) the former effect mainly because the ratio of human capital in local jobs to the one in national jobs, $\frac{H_l}{H_n}$, is low (high) and thus the marginal effect of increased $\frac{H_l}{H_n}$ on w_n is large (small) [see (2)].[23]

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 $a(= e)$ and in (a) and (b). First, consider case (a) $\bar{e} \leq \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$. When the level of wealth is high enough that $a = e > \alpha\left(\frac{h_l}{\delta_l} + \bar{e}\right)$, earnings increase (decrease) with s_l for $s_l < (>) s_l^\star(e)$ and are highest 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. As $e(= a)$ increases, the graph shifts upward and $s_l^\star(e)$ increases. By contrast, when $a = e \leq \alpha\left(\frac{h_l}{\delta_l} + \bar{e}\right)$, earnings *decrease with* s_l *for any* s_l and thus are highest at $s_l = 0$. That is, although higher s_l means a higher proportion of expenditure allocated to the education of the skill for local jobs, *no allocation* to the education maximizes earnings of workers with little wealth, who choose a local job.

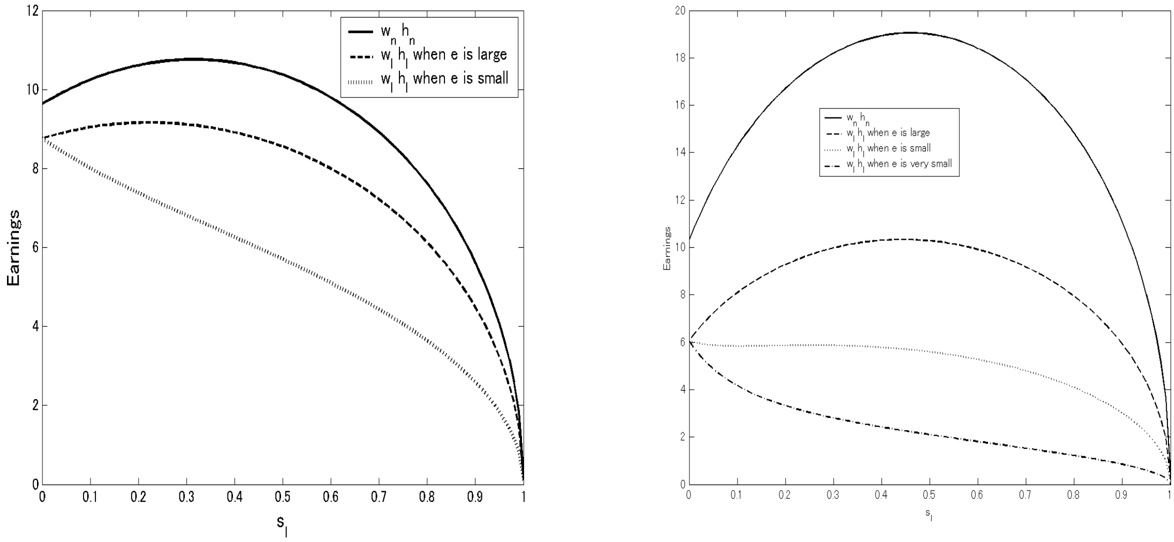
An increase in s_l raises human capital $h_l(e, s_l)$ but lowers wage rate w_l of workers with a local job. When s_l is low (high), the former effect tends to dominate (be dominated by) the latter effect, mainly because $\frac{H_n}{H_l}$ is relatively high (low) and thus the marginal effect of decreased $\frac{H_n}{H_l}$ on w_l is small (large). Further, the positive effect through human capital *increases with* e, because one with greater wealth and thus educational spending benefits more from the increased weight on the education of the skill useful for local jobs. Hence, earnings of a worker with a local job increase (decrease) with s_l for small (large) s_l and earning-maximizing 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 is dominated by the negative effect through the wage rate even at $s_l = 0$, thus earnings are highest at $s_l = 0$.

In case (b) $\bar{e} > \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$, the result is similar to case (a) for large a and for very small a, but, for the intermediate range of a, earnings of workers with a local job decrease with s_l for $s_l < s_l^\circ(e)(s_l^{\circ\prime}(e) < 0)$, 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.

Finally, the case in which the return to education for local jobs is positive and $e^+ < \bar{e}$ holds (the upper-middle region of Figure 3) is considered. Here, only the main result is presented and the details are relegated to Section A.1 of Appendix A. Unlike the previous cases, analytical results *cannot* be obtained for some ranges of s_l and as for earnings of workers with a local job, $a(= e)$. However, the analytical result in the appendix and numerical simulations suggest that results for workers with a local job are qualitatively the same as the case $e^+ = \bar{e}$. Results for workers with a national job too are qualitatively unchanged *unless* the proportion of those with limited wealth is very high, in which case their earnings decrease with s_l for any s_l.[24] Figure 6 presents a numerical

[23] Assuming a positive effect of the education of the skill for local jobs on the development of the skill for national jobs would not affect the result qualitatively. As explained in footnote 19, when s_l is large, it is almost certain that the effect of s_l on h_n remains negative and thus the result does not change. When s_l is small, an increase in s_l raises h_n if the positive indirect effect of s_l on h_n through the education of the skill for local jobs outweighs the negative direct effect. Even if this is the case, it is almost certain that the effect of s_l on $\frac{H_l}{H_n}$ and thus w_n remains positive or is small negative, because an increase in s_l also raises h_l. Hence, the result would not change when s_l is small also.

[24] In this case, numerical simulations suggest that $w_l h_l$ for $a = e \leq e^+$ (i.e., earnings of those who *actually* choose a local job) too decreases with s_l. These results indicate that earnings of *all workers* decrease with s_l when a large



(a) When $\overline{e} \leq \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$ (b) When $\overline{e} > \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$

Figure 6: Effect of s_l on earnings when the return is positive and $e^+ < \overline{e}$

example of the relationship between s_l and earnings of workers with a local job when $a(= e)$ is large, small, and very small (in (b) only), and of workers with a national job.[25]

5 Main Results

The previous section examines the effects of s_l on earnings for the three cases separately by assuming that the economy is in a particular case for any s_l. Proposition 1 (Figure 3) in Section 3, however, shows that which of the cases is realized depends on s_l and other parameters. By taking into account Proposition 1 as well as the results in the previous section, this section analyzes the effects of s_l on earnings *net of educational spending*. The 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 (thus the value of s_l).

5.1 When $F(\overline{e}) \leq 1 - \alpha$

proportion of people have limited wealth for education.

[25]In both (a) and (b), $\alpha = 0.5$, $\underline{h}_l = \delta_n/\delta_l = 0.5$, the distribution of wealth follows truncated log normal with maximum 100, mean 6 and variance 10, and $a = e = \overline{e}$ for the earning profile for national jobs. In (a), $\delta_l = 0.25$, $\overline{e} = 6$, $A = 30$, and the value of e of the profile for local jobs is $1.1\alpha\left(\frac{h_l}{\delta_l} + \overline{e}\right) = 4.4$ when e is large and $0.4\Lambda(\overline{e}) = 1.6$ when e is small, where $\Lambda(\overline{e}) \equiv \frac{\alpha\left(\frac{h_l}{\delta_l}+\overline{e}\right)}{1+\frac{h_l}{4\underline{h}\overline{e}}\left[(1-\alpha)\overline{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]^2}$, while in (b), $\delta_l = 1$, $\overline{e} = 10$, $A = 10$, and values of e of the profile when e is large, small, and very small are respectively $\alpha\left(\frac{h_l}{\delta_l} + \overline{e}\right) = 5.25$, $0.87\Lambda(\overline{e}) = 2.4$, and $0.1\Lambda(\overline{e}) = 0.2759$.

The next proposition examines the effects of s_l on net earnings when the share of individuals facing the wealth constraint on educational investment is low enough that $F(\overline{e}) \leq 1 - \alpha$ holds, in which case $e^+ = \overline{e}$ always holds (see Figure 3). Appendix B presents a fully detailed statement of the proposition.

Proposition 4.[26] *Suppose that $F(\overline{e}) \leq 1 - \alpha$ and thus $e^+ = \overline{e}$ hold.*

(*i*) *If A, δ_n, or δ_l is small enough that $\underline{s_l} > s_l^{\star\star}$, net earnings of all workers decrease with s_l.*[27]

(*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, δ_n, and δ_l are sufficiently large.*

(*b*) *Net earnings of workers 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 workers with wealth below the threshold decrease with s_l for any s_l. Net earnings of the former workers with wealth $a = e$ are maximized at $s_l = s_l^{\star}(e)$, when A, δ_n, and δ_l are sufficiently large.*

(*c*) *$\underline{s_l}$ in (a), $\max\{\underline{s_l}, s_l^{\circ}(e)\}$ and the threshold wealth in (b) decrease with A, δ_n, and δ_l.*

The first part of the proposition shows that, if A, δ_n, or δ_l is small, net earnings of *all workers* decrease with s_l *for any* s_l. This implies that allocating educational spending completely to the education of the skill for national jobs, i.e., lingua franca education, maximizes net earnings of everyone, when TFP is low or education is not effective in skill development.

The result can be explained as follows. Proposition 2 in the previous section shows that, when s_l is low or high enough that the return to education for local jobs is negative, i.e., $s_l < \underline{s_l}$ or $s_l > \overline{s_l}$, net earnings of everyone decrease with s_l. Proposition 3 shows that, when s_l is in the intermediate range and thus the return is positive, net earnings of those with wealth below a threshold decrease with s_l, while net earnings of wealthier individuals increase (decrease) with small (large) s_l and are highest at intermediate s_l, i.e., $s_l^{\star\star}$ or $s_l^{\star}(e)$. When TFP is low or education is not effective, the return to education is low *for given* s_l. Hence, the return to education for local jobs is negative for a wide range of s_l, i.e., $\underline{s_l}$ is large and $\overline{s_l}$ is small. Thus, the range of s_l for which the return is positive and net earnings of the wealthier increase with s_l becomes *ineffective*, i.e., $\underline{s_l} > s_l^{\star\star}(> s_l^{\star}(e))$.[28]

[26] As used in Figure 3, $\underline{s_l}$ ($\overline{s_l}$) is the lower (higher) critical level of s_l at which the return to education for workers with a local job is 0 when $e^+ = \overline{e}$.

[27] $\underline{s_l}$ decreases with A, δ_n, and δ_l from Proposition 1 (i)(b) and $s_l^{\star\star} \equiv (1 - \alpha) - \alpha \frac{h_l}{\delta_l \overline{e}}$.

[28] The most important and perhaps surprising results of the paper would be (i) and the result for individuals with little wealth of (ii)(b) of the proposition, and the corresponding results for the case $e^+ < \overline{e}$ below (Proposition 6). Assuming a positive effect of mother tongue education [education of the skill for local jobs] on the development of the lingua franca skill [the skill for national jobs] (footnote 9) would hardly affect the results. The result for those with little wealth is clearly not affected, because they choose a local job. As the explanation of the result in the main text suggests, (i) of the proposition relies on Propositions 1–3, in particular, the results that the return to education for local jobs is negative (positive) for $s_l < \underline{s_l}$ and $s_l > \overline{s_l}$ (for $s_l \in (\underline{s_l}, \overline{s_l})$), net earnings decrease with s_l when the return is negative and when the return is positive and s_l is large. Footnotes 14, 19, and 23 show that assuming

By contrast, if A, δ_n, and δ_l are not small, net earnings of those with wealth above a certain level, who choose a national job or a local job depending on 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 5 (Proposition 3). Either an intermediate level of s_l ($s_l^{\star\star}$ for those with a national job and $s_l^{\star}(e)$ for those with a local job and $a = e$) or $s_l = 0$ maximizes their net earnings, and when TFP and the effectivness of education are sufficiently high, allocating the expenditure to both types of education is economically optimal for them.

Finally, net earnings of those with wealth below the threshold, who choose a local job, decrease with s_l *for any* s_l. Hence, their net earnings are maximized at $s_l = 0$ *regardless of* the level of TFP and the effectiveness of education. A greater emphasis on the education useful for their jobs lowers the earnings, because a positive effect of higher s_l on their human capital is small due to their limited educational spending and is dominated by a negative effect on the wage rate of local jobs due to decreased human capital in complementary national jobs. What is crucial for this result is the assumption that human capital for local jobs, i.e., the mother tongue skill, is positive without education, i.e., $\underline{h}_l > 0$. If $\underline{h}_l = 0$, net earnings of *everyone* increase (decrease) with s_l for $s_l < (>) 1 - \alpha$ when the return to education for a local job is positive. The assumption makes the negative effect of higher s_l on earnings through decreased w_l greater relative to the positive effect through increased h_l for those with very small wealth.

The last part of the proposition shows that as A, δ_n, and δ_l increase, the wealth threshold falls and the range of s_l over which net earnings of those with wealth above the threshold increase with s_l expands. The result implies that higher TFP and more effective education make a higher proportion of people benefit from the balanced education.

The first result implies that, if TFP or the effectiveness of education is low, dual education lowers net earnings of workers *irrespective of their wealth.* This can be seen clearly by considering the case in which everyone has enough wealth to make optimal educational investment, i.e., $F(\bar{e}) = 0$.

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

(*i*) *If* A, δ_n, *or* δ_l *is small enough that* $\underline{s_l} > s_l^{\star\star}$, *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, δ_n, *or* δ_l *are sufficiently large.*

Without the wealth constraint, if A, δ_n, or δ_l are large enough, a balanced allocation of expenditure to both types of education ($s_l = s_l^{\star\star}$) maximizes net earnings of everyone, but if not, the full allocation to the education of the skill for national jobs remains economically optimal for all.

the complementarity in skill development would not affect these results qualitatively, except that net earnings of workers with a national job might increase with s_l if $s_l < \underline{s_l}$ and the positive indirect effect of s_l on h_n through the education of the skill for local jobs is large and outweighs the negative direct effect. If this is the case, their net earnings increase with s_l for very small s_l and decrease with s_l for greater s_l when A, δ_n, or δ_l is small.

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

When the proportion of those whose educational investment is constrained by wealth 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 4 applies. When $e^+ < \overline{e}$, Proposition 6 of Appendix A presents analytical results. The results are mostly similar to the case $e^+ = \overline{e}$, although unlike before, the proposition does not cover intermediate ranges of A, δ_n, δ_l, and a and as for one result, some ranges of s_l.[29] As when $e^+ = \overline{e}$, the full allocation of expenditure to the education of the skill for national jobs, i.e., lingua franca education, is economically optimal for those with little wealth, and when TFP or the effectiveness of education in skill development is low, for everyone; while when TFP or the effectiveness of education is high, the balanced education maximizes net earnings of those with sufficient wealth.

When the proportion of wealth-constrained individuals is very high, however, the result is qualitatively different from the case $e^+ = \overline{e}$. In this case, an analytical result for those with a national job (Proposition 5 (i)(a) of Appendix A) and numerical simulations for those with a local job suggest that net earnings of everyone are highest under lingua-franca-only education.

5.3 Policy and other implications

This section discusses policy and other implications of the above results.

5.3.1 Conflict between educational and economic outcomes of bilingual education

As mentioned in Section 3.3, a general consensus among specialists on language and education is that using mother tongues at least in primary education is effective for students to acquire adequate skills (Heugh, 2011a), and empirical studies in economics (Angrist, Chin and Goody, 2008; Ramachandran, 2017) find that the introduction of mother tongue education has a positive effect on academic skills and education. Consistent with them, Proposition 1 implies that balanced education of lingua franca and mother tongue skills (moderate s_l) is critical for skill development of those who choose a local job due to limited wealth.

The results in the previous sections, however, show that the advantage of the balanced education in skill development *does not* necessarily translate into high *economic* returns of such education. Earnings net of educational spending of those with little wealth decrease with s_l and thus lingua-franca-only education brings the highest return to them.[30] Further, when educational and technological conditions of a country are poor, i.e., the level of TFP is low or the effectiveness of education in skill development is low, net earnings of everyone decrease with s_l and such

[29] Unlike the case $e^+ = \overline{e}$, net earnings change discontinuously when the return to educational investment for local jobs turns from negative to positive or the other way around with a change in s_l. This is because $\frac{H_n}{H_l}$ changes discontinuously with the discontinuous change in educational spending by the poor. When $e^+ = \overline{e}$ (Proposition 4), by contrast, net earnings change continuously because a discontinuous change in π_n (the proportion of workers with a national job among those with $a \geq e^+ = \overline{e}$) makes the change of $\frac{H_n}{H_l}$ continuous.

[30] To be precise, when $e^+ < \overline{e}$, their net earnings decrease with s_l except at the higher s_l such that the return to education for local jobs is 0. The net earnings, however, are always highest at $s_l = 0$. A similar caveat applies to the next sentence as well.

education is best for all in terms of the economic outcome. In the real economy, these conditions tend to be related to the level of economic and social development of a country. Hence, the result suggests that if the level of development is reasonably high, the balanced education is economically desirable except for the very poor; otherwise, lingua-franca-only education is desirable in terms of the economic outcome. Consistent with this implication, Angrist and Lavy (1997) find a significant increase in the weight on mother tongue education lowers wages in a developing country, Morocco.[31]

These results imply that a *trade-off between educational and economic outcomes* exists for those with little wealth and when the educational and technological conditions are poor, for all: under lingua-franca-only (balanced bilingual) education, their net earnings are highest (low) but their mother tongue skill, which is an essential skill in daily life even for those with a national job, is lowest (high).

The results also imply that improved academic performance of students after the expansion of mother tongue 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 due to very low s_l, the government can turn the return to positive by raising s_l appropriately, and can boost educational investment (from 0 to a) of those who have limited wealth and end up in a local job. The policy change succeeds in raising their skill. However, it always *lowers* net earnings of the very poor (it could raise their *gross* earnings), and when educational and technological conditions of a country are not good, lowers net earnings of others also.

5.3.2 Socially desirable policies

Then, what kind of policies should be implemented when both educational and economic outcomes are taken into account? Suppose that the government chooses the education policy s_l along with a redistributive policy (tuition subsidy or income transfer) financed by lump-sum tax to maximize a social welfare function that depends on both the educational outcome (the mother tongue skill h_l) and the economic outcome (consumption, which equals the sum of net earnings, wealth, and, if any, means-tested transfer minus tax) of individuals.[32]

The above results suggest that, when educational and technological conditions of a country are good (i.e, TFP is reasonably high and education is reasonably effective), the welfare-maximizing government implements balanced bilingual education *together with* a redistributive policy that enables those with little wealth to expend sufficiently on education.[33] Redistribution toward the

[31] Angrist and Lavy (1997) find that replacing French (the common language in the modern sector of the economy even after the reform) with Arabic as the medium of instruction in post-primary education greatly lowered returns to schooling in Morocco.

[32] The educational outcome is measured by the mother tongue skill because, as mentioned above, it is an essential skill in daily life even for individuals with a national job.

[33] The mother tongue skill $h_l = \underline{h}_l + \delta_l s_l e$ is maximized at s_l very slightly less than $\overline{s_l}$ or $\overline{s_l}(f)$ ($\overline{s_l}(f)$ denotes the higher s_l such that the return to education for local jobs is 0 when $e^+ < \overline{e}$) for those choosing a local job from Proposition 1 (i)(a) and at $s_l = 1$ for those choosing a national job. *Without redistribution*, consumption is maximized at $s_l = 0$ for the very poor and at some $s_l \in (\underline{s_l}, \overline{s_l})$ or $(\underline{s_l}(f), \overline{s_l}(f))$, whose value differs depending on e and thus a, for others from Proposition 4 (ii) and Proposition 6 (ii) and (iii) in Appendix A. A redistributive policy

wealth-constrained very poor is socially desirable because the policy not only raises their consumption significantly but also makes it higher under the balanced education, which also brings good economic outcomes to the wealthier and good educational outcomes to all.

By contrast, when the educational and technological conditions are not good, the policy tools cannot bring good educational and economic outcomes to everyone: the trade-off between the two outcomes exists for all and as Corollary 1 shows, redistributive policies cannot change it. Then, the government that balances the educational outcome against the economic one would choose bilingual education with a *smaller* (*but not too small*) weight on mother tongue education than under the more favorable conditions, e.g., s_l slightly greater than $\underline{s_l}$ when $e^+ = \bar{e}$, together with redistribution toward the very poor.[34] Such a policy achieves the better educational outcome than lingua-franca-only education at the relatively small loss of consumption.

Proposition 2 also implies that bilingual education with a *very small* weight on mother tongue education, e.g., $s_l < \underline{s_l}$ when $e^+ = \bar{e}$, is *worse* than lingua-franca-only education: net earnings of everyone are lower than and mother tongue skills of those who have limited wealth and thus choose a local job are as low as under the latter education.

Note that the model does not take into account possibly important effects of the choice of languages in education. Mother tongue education would raise the ethnic language skill and contribute to the accumulation of social capital in the local ethnic community. It might also stimulate political participation and increase support for democracy (Albaugh, 2016). Lingua franca education, on the other hand, would help people identify with the nation and contribute to national unity and stability in an ethnically diverse society. It might also reduce linguistic diversity and promote public goods provision and economic growth (Desmet, Ortuño-Ortín, and Wacziarg, 2012). Policy implementation in the actual society needs to take into account these effects as well as the skill and earnings effects considered in the model.

Finally, the result that the redistributive policies are essential for the very poor to benefit economically from the balanced education gives another justification for governmental support of basic education, in addition to usual rationales based on positive externality, human rights, and so on, in multilingual countries.

that, at the expense of the rich, induces the wealth-constrained very poor to spend sufficiently more on education not only raises their consumption significantly for given s_l (the increase is greater than the magnitude of consumption decrease of the unconstrained rich) but also makes it highest at an intermediate s_l. Hence, the education policy with some $s_l \in (\underline{s_l}, \overline{s_l})$ or $(\underline{s_l}(f), \overline{s_l}(f))$ together with such a redistributive policy would maximize social welfare.

[34]As mentioned in footnote 33, the mother tongue skill h_l is highest at s_l very slightly less than $\overline{s_l}$ or $\overline{s_l}(f)$ for those choosing a local job, i.e., the poor, and at $s_l = 1$ for those choosing a national job. At the same time, unlike when the conditions are good, consumption of everyone decreases with s_l and is highest at $s_l = 0$. Hence, the welfare-maximizing s_l is smaller than under the better conditions. Further, unless the weight on the educational outcome and the one on the poor in the social welfare function are small, social welfare would be higher when $e > 0$ holds for the poor and s_l is relatively close to 0, e.g., s_l slightly greater than $\underline{s_l}$ ($\underline{s_l}(f)$) when $e^+ = (<)\bar{e}$. Redistribution toward the very poor raises social welfare because as mentioned in footnote 33, an increase in their consumption is greater than the magnitude of consumption decrease of the rich.

6 Conclusion

In many developing countries, people use their mother tongue in daily life and in local business, but use the language of the former colonizer as the lingua franca in national business and in communications with other ethnic groups. How much weight should be placed on teaching one's mother tongue and teaching the lingua franca and which language should be used as a language of instruction of other subjects are critical issues in these countries.

Specialists on language and education generally stress the importance of mother tongue education, at least in primary education, for skill development. By contrast, we know very little what is a desirable combination of mother tongue and lingua franca education in terms of earnings and what kind of policies should be conducted when both educational and economic outcomes are taken into account.

This paper has developed a simple model to examine these issues. It is shown that balanced education of the two languages is critical for skill development of students with limited wealth for education. It is also found that balanced bilingual education yields higher earnings net of educational expenditure than lingua-franca-only education only when the country has favorable educational and technological conditions, i.e., productivity is reasonably high and education is reasonably effective in skill development, and only for those with sufficient wealth. This implies that a trade-off between educational and economic outcomes exists for those with little wealth and under unfavorable conditions, for everyone.

Main policy implications are as follows. When educational and technological conditions of a country are favorable, the government that takes into account both educational and economic outcomes of individuals would implement balanced bilingual education together with a redistributive policy that enables those with little wealth to expend sufficiently on education. By contrast, when the conditions are not good, the welfare-maximizing government would choose bilingual education with a smaller (but not too small) weight on mother tongue education than under the more favorable conditions (together with redistribution toward the very poor).

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] Albaugh, Ericka A. (2016), "Language, education, and citizenship in Africa", Afrobarometer WP No. 162.

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

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

[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] Baker, Colin and Wayne E. Wright (2017), *Foundations of Bilingual Education and Bilingualism*, 6th edition, Multilingual Matters.

[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] Cappellari, Lorenzo and Antonio Di Paolo (2018), "Bilingual schooling and earnings: evidence from a language-in-education reform," *Economics of Education Review* 64, 90−101.

[9] Chakraborty, Tanika and Shilpi Kapur Bakshi (2016), "English language premium: Evidence from a policy experiment in India," *Economics of Education Review* 50 (1), 1−16.

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

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

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

[13] 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.

[14] 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.

[15] 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.

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

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

[18] 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.

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

[20] Ramachandran, Rajesh (2017), "Language use in education and human capital formation: evidence from the Ethiopian educational reform," *World Development* 98, 195−213.

[21] 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.

[22] Yuki, Kazuhiro (2021), "Language education and economic outcomes in a bilingual society," MPRA paper 106119.

Appendix A: Analysis of the effect of s_l on earnings when $e^+ < \overline{e}$

This appendix examines in detail the effect of s_l on earnings when $e^+ < \overline{e}$.

A.1 When the return to education for local jobs is constrained to be positive

First, as in Section 4.2, the effect is examined under the assumption that the return to educational investment for local jobs is positive for any s_l. The next proposition summarizes analytical results, based on Proposition A1 in Appendix B, which presents a fully detailed statement. This proposition is the counterpart of Proposition 3 when $e^+ = \overline{e}$. 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 5. *Suppose that the return to education for local jobs is positive and* $e^+ < \overline{e}$ *holds.*

(i) (a) If the proportion of those with limited wealth is high enough that $e^+(0) \leq \frac{\alpha}{1-\alpha}\frac{h_l}{\delta_l}$ *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 *maximizing* $w_n h_n$ *is smaller than* $s_l^{\star\star}$*, the critical* s_l *when* $e^+ = \overline{e}$.

(ii) (a) If $\overline{e} \leq \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$, $\frac{d(w_l h_l)}{ds_l} > (<)\ 0$ *for small (large)* s_l *when* $a(= e)$ *is large, while* $\frac{d(w_l h_l)}{ds_l} < 0$ *for any* s_l *when* $a(= e)$ *is small. In the former case,* s_l *maximizing* $w_l h_l$ *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* $a(= e)$ *is intermediate, while when it is small and large, results are similar to (a). In the former case,* s_l *maximizing (minimizing)* $w_l h_l$ *is greater (smaller) than* $s_l^{\star}(e)$ $(s_l^{\circ}(e))$*, the critical* s_l *when* $e^+ = \overline{e}$.

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

As shown in Proposition A1 in Appendix B, unlike the case $e^+ = \overline{e}$, analytical results *cannot* be obtained for some ranges of s_l and as for $w_l h_l$, $a(= e)$. However, the above proposition and numerical simulations suggest that results for workers with a local job are qualitatively the same as the case $e^+ = \overline{e}$: when $\overline{e} \leq \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$, $w_l h_l$ increases (decreases) with s_l for small (large) s_l when $a(= e)$ is large, and $w_l h_l$ decreases with s_l for any s_l when it is small; when $\overline{e} > \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$, $w_l h_l$ 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 $a(=e)$ is intermediate, while the results when it is large and small are similar to the previous case. As for the relationship between s_l maximizing $w_l h_l$ and $a(=e)$, unlike the case $e^+ = \overline{e}$, an analytical result is not obtained but numerical simulations suggest that the relationship is positive as before.

There exist minor differences from the case $e^+ = \overline{e}$. First, s_l maximizing earnings of workers with a local job and a above the threshold is greater than $s_l^\star(e)$, the critical s_l when $e^+ = \overline{e}$. Second, the maximum level of $a(=e)$ 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, for given s_l, $e^+ < (=)\overline{e}$ holds when $F(\overline{e})$ is relatively large (small). Hence, these results suggest that individuals choosing a local job are more likely to benefit from the education of the skill useful in their future jobs when the share of those who face the wealth constraint on educational investment, 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 are also similar to the case $e^+ = \overline{e}$, but some differences exist. First, if the proportion of those with limited wealth is high enough that $e^+(0) \leq \frac{\alpha}{1-\alpha}\frac{h_l}{\delta_l}$, $w_n h_n$ decreases with s_l *for any* s_l. Numerical simulations suggest that $w_l h_l$ of those who *actually* choose a local job (i.e., those with $a \leq e^+(s_l)$) also decreases with s_l in this case. These results indicate that earnings of *all workers* decrease with s_l when a large proportion of people have limited wealth for education. Second, when $e^+(0) > \frac{\alpha}{1-\alpha}\frac{h_l}{\delta_l}$, s_l maximizing $w_n h_n$ is smaller than $s_l^{\star\star}$ when $e^+ = \overline{e}$. Hence, in contrast to workers with a local job, workers who have abundant wealth to choose a national job are *less* likely to benefit from the education of the skill for local jobs when $e^+ < \overline{e}$ than when $e^+ = \overline{e}$.

A.2 When the return to education for local jobs is endogenous

Now, as in Section 5, the effects of s_l on earnings net of educational spending is examined by taking into account the fact that whether the return to education for local jobs is positive or not is endogenously determined, drawing on Propositions 1 and 2 as well as Proposition 5. The following proposition is the counterpart of Proposition 4 when $e^+ = \overline{e}$. Note that unlike when $e^+ = \overline{e}$, the proposition does not cover intermediate ranges of A, δ_n, δ_l, and a and as for (iii) of the proposition, some ranges of s_l. In what follows, $\underline{s_l}(f)$ ($\overline{s_l}(f)$) denotes the lower (higher) s_l such that the return to education for local jobs is 0 when $e^+ < \overline{e}$. ("f" is to indicate the dependence of their values on the distribution of wealth.) Appendix B provides a fully detailed statement of the proposition.

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

(i) When A, δ_n, or δ_l is small, net earnings of all workers decrease with s_l except at $s_l = \underline{s_l}(f)$, $s_l = \overline{s_l}(f)$, or both, where they increase discontinuously.[35] *The net earnings are maximized at $s_l = 0$ if A, δ_n, or δ_l is sufficiently small.*

[35] 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 either $s_l = \underline{s_l}(f)$, $s_l = \overline{s_l}(f)$, or both.

(*ii*) *Those with wealth below a certain level choose a local job for any* s_l*. 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*) *When* A*,* δ_n*, and* δ_l *are not small, 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*. When* δ_l *is sufficiently large, their net earnings are maximized at* $s_l \in (\underline{s_l}(f), \overline{s_l}(f))$*.*[36]

The first part of the proposition shows that, if A, δ_n, or δ_l is small, net earnings of everyone decrease with s_l except at either $s_l = \underline{s_l}(f)$, $s_l = \overline{s_l}(f)$, or both, where, differently from the case $e^+ = \bar{e}$, they increase discontinuously. But, as before, net earnings are highest at $s_l = 0$ if A, δ_n, or δ_l is sufficiently small. The second part shows that irrespective of values of A, δ_n, and δ_l, net earnings of those with little wealth decrease with s_l except at $s_l = \overline{s_l}(f)$, where, unlike when $e^+ = \bar{e}$, the earnings increase discontinuously, but as before, are highest at $s_l = 0$. The last part shows that, when A and δ_n are not small and δ_l is sufficiently large, net earnings of those with sufficient wealth are highest at $s_l \in (\underline{s_l}(f), \overline{s_l}(f))$.

Appendix B: More precise statements for several propositions and a claim

This appendix presents more precise statements for Proposition 1 in Section 3.3, Proposition 4 in Section 5, and Propositions 5 and 6 in Appendix A. The appendix also presents Claim 1 that is used for drawing Figure 3 in Section 2.

Proposition 1. *Suppose that* A *is not extremely low so that* $w_l\delta_l s_l - 1 = (1-\alpha)A\left[\frac{\alpha\delta_n(1-s_l)e^+}{(1-\alpha)(\underline{h}_l+\delta_l s_l e^+)}\right]^\alpha \delta_l s_l - 1 > 0$ *at* s_l *satisfying* $\underline{h}_l\left[(1-s_l-\alpha s_l)\underline{h}_l F(e^+) + (1-s_l-\alpha)\delta_l s_l \int_0^{e^+} e f(e)de\right] + \frac{1}{\alpha}[(1-s_l-\alpha s_l)\underline{h}_l + (1-s_l-\alpha)s_l\delta_l e^+] \times h_l(e^+, s_l)e^+ f(e^+) = 0$ *when* $e^+ < \bar{e}$*.*[37] *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 individuals 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*), δ_n *and* δ_l (*respectively, effectiveness of education of the skill for national jobs and for local jobs*).

(*ii*) *When* s_l *is small or large enough that the return to educational investment for local jobs is negative,* $e^+ < (=) \bar{e}$ *holds iff* $F(\bar{e})$ *is large* (*small*) *enough that* $\frac{F(\bar{e})-(1-\alpha)}{[F(\bar{e})]^\alpha[1-F(\bar{e})]^{1-\alpha}} A(\delta_n(1-s_l))^\alpha \left(\frac{\underline{h}_l}{\bar{e}}\right)^{1-\alpha} > (\leq) 1$*. When the return is positive,* $e^+ < (=) \bar{e}$ *holds iff* $F(\bar{e}) > (\leq) 1 - \alpha\frac{\underline{h}_l + \delta_l s_l \int_0^{\bar{e}} e f(e)de}{\underline{h}_l + \delta_l s_l (1-\alpha)\bar{e}}$*. When* $F(\bar{e}) \leq 1-\alpha$*,* $e^+ = \bar{e}$ *always holds.*

[36] 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 the directions of change.

[37] e^+ does not depend on A.

Proposition 4 . *Suppose that* $F(\bar{e}) \leq 1-\alpha$ *and thus* $e^{+} = \bar{e}$ *hold.*

(i) If A, δ_n, *or* δ_l *is small enough that* $\underline{s_l} > s_l^{\star\star}$, *net earnings of all workers decrease with* s_l.[38]

(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, δ_n, *and* δ_l *are large enough that* $\left[\frac{(1-\alpha)^2}{\alpha^2}\right]^{1-\alpha}\alpha\left(1+\frac{h_l}{\delta_l\bar{e}}\right)(\delta_l)^{1-\alpha} > \left(\frac{\{F(\bar{e})+(1-\pi_n)[1-F(\bar{e})]\}\underline{h_l}}{\pi_n[1-F(\bar{e})]\bar{e}}\right)^{1-\alpha}$.[39]

(b) Net earnings of workers 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 workers with wealth below the threshold decrease with* s_l *for any* s_l. *The threshold equals* e *satisfying* $s_l^{\star}(e) = \underline{s_l}$ *if* $\bar{e} \leq \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$ *and if* $\bar{e} > \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$ *and* $s_l^{\star}(\Lambda(\bar{e})) \leq \underline{s_l}$, *where* $\Lambda(\bar{e}) \equiv \frac{\alpha\left(\frac{h_l}{\delta_l}+\bar{e}\right)}{1+\frac{\delta_l}{4\underline{h}\bar{e}}\left[(1-\alpha)\bar{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]^2}$, *otherwise, it equals* $\Lambda(\bar{e})$. *Net earnings of the former workers with wealth* $a = e$ *are maximized at* $s_l = s_l^{\star}(e)$, *when* A, δ_n, *and* δ_l *are large enough that* $A(\delta_n\bar{e})^{\alpha}\left\{(1-\alpha)\left[\frac{\alpha(1-s_l^{\star}(e))}{1-\alpha}\right]^{\alpha}(\underline{h}_l+\delta_l s_l^{\star}(e)\bar{e})^{1-\alpha}-\alpha\left(\frac{\{F(\bar{e})+(1-\pi_n)[1-F(\bar{e})]\}\underline{h_l}}{\pi_n[1-F(\bar{e})]}\right)^{1-\alpha}\right\}+\bar{e}-e > 0.$

(c) $\underline{s_l}$ *in (a),* $\max\{\underline{s_l}, s_l^{\circ}(e)\}$ *and the threshold wealth in (b) decrease with* A, δ_n, *and* δ_l.

Proposition 6 . *Suppose that* $e^{+} < \bar{e}$ *holds.*

(i) When A, δ_n, *or* δ_l *is small enough that either* $e^{+}(0) \leq \frac{\alpha}{1-\alpha}\frac{h_l}{\delta_l}$ *or* $\underline{s_l}(f) \geq s_l^{\sharp}$ *(*$s_l^{\sharp}$ *satisfies* $s_l^{\sharp} = (1-\alpha)-\alpha\frac{h_l}{\delta_l e^{+}(s_l^{\sharp})}$*) and* $\underline{s_l}(f) \geq s_{l,h}^{\bigtriangledown}(e)$ *hold,*[40] *net earnings of all workers decrease with* s_l *except at* $s_l = \underline{s_l}(f)$, $s_l = \overline{s_l}(f)$, *or both, where they increase discontinuously.*[41] *The net earnings are maximized at* $s_l = 0$ *if* A, δ_n, *or* δ_l *is small enough that* $\left(\frac{1-\alpha}{\alpha}\frac{h_l}{\delta_n e^{+}(0)}\right)^{1-\alpha} > \left[(1-\alpha)^2 A\delta_l\right]^{\frac{1-\alpha}{\alpha}}\alpha.$

(ii) Those with wealth below a certain level choose a local job for any s_l. *The threshold wealth equals* $\min\left\{e^{+}(0), \alpha\left[\frac{h_l}{\delta_l}+E(e|e<e^{+}(0))\right]\right\}$ *when* $E(e|e<e^{+}(0)) \leq \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$ *and equals* $\min\{e^{+}(0), \Omega(e^{+}(0))\}$ *when* $E(e|e<e^{+}(0)) > \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$, *where* $\Omega(e^{+}(0)) \equiv \frac{\alpha\left(\frac{h_l}{\delta_l}+E(e|e<e^{+}(0))\right)}{1+\frac{\delta_l}{4\underline{h}_l E(e|e<e^{+}(0))}\left[(1-\alpha)E(e|e<e^{+}(0))-(1+\alpha)\frac{h_l}{\delta_l}\right]^2}$. *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) When A, δ_n, *and* δ_l *are large enough that* $e^{+}(0) > \frac{\alpha}{1-\alpha}\frac{h_l}{\delta_l}$, $E(e|e<e^{+}(0)) > \alpha\max\left\{\frac{1}{1-\alpha}\frac{h_l}{\delta_l}, \frac{e^{+}(0)}{1+\alpha}\right\}$, *and* $\min\{s_l^{\flat}, s_{l,h}^{\bigtriangleup}(e)\} > \underline{s_l}(f)$ *(*$s_l^{\flat}$ *satisfies* $s_l^{\flat} = (1-\alpha)-\alpha\frac{h_l}{\delta_l E(e|e<e^{+}(s_l^{\flat}))}$*), net earnings of those*

---

[38] $\underline{s_l}$ decreases with A, δ_n, and δ_l from Proposition 1 (i)(b) and $s_l^{\star\star} \equiv (1-\alpha)-\alpha\frac{h_l}{\delta_l\bar{e}}$.

[39] The LHS increases with δ_l and the RHS decreases with A and δ_n because π_n increases with them.

[40] Note that $e^{+}(0)$ does not depend on A, δ_n, and δ_l, $\underline{s_l}(f)$ decreases with these variables, and $s_l^{\sharp}$ increases with δ_l. $s_{l,h}^{\bigtriangledown}(e)$ is defined in Proposition A1 (ii)(a) and $\underline{s_l}(f) \geq s_{l,h}^{\bigtriangledown}(e)$ holds when A, δ_n, and δ_l are small.

[41] Net earnings of workers who have abundant wealth and thus choose a national job for any s_l, i.e., $a \geq \min\{e^{+}(1), \bar{e}\}$, 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 either $s_l = \underline{s_l}(f)$, $s_l = \overline{s_l}(f)$, or both.

with wealth greater than $\max\left\{\alpha\left(\frac{h_l}{\delta_l}+e^+(0)\right), \Lambda(\overline{e})\right\}$ *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*.*[42] *When* δ_l *is large enough that* $\left(\frac{h_l}{e^+(0)}\right)^{1-\alpha} < \left(\frac{h_l+\delta_l s_l^\flat \overline{e}}{\overline{e}}\right)^{1-\alpha}(1-s_l^\flat)^\alpha$ *for those with a national job and* $(1-\alpha)A\left(\delta_n\frac{\alpha}{1-\alpha}\right)^\alpha\{(\alpha e)^\alpha[\underline{h}_l+\delta_l(1-\alpha)e]^{1-\alpha} -(e^+(0))^\alpha(\underline{h}_l)^{1-\alpha}\}-e > 0$ *for those with a local job and wealth* $a = e$*, their net earnings are maximized at* $s_l \in (\underline{s_l}(f), \overline{s_l}(f))$*.*[43,44]

The following proposition presents more precise results for Proposition 5. Note that except (i)(a) of the proposition, conditions are sufficient but not necessary. Proofs of the proposition and the claim below are contained in Appendix D posted on the author's webpage (http://www.econ.kyoto-u.ac.jp/˜yuki/english.html).

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

(i) (a) If $e^+(0) \le \frac{\alpha}{1-\alpha}\frac{h_l}{\delta_l}$*, which is true when 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* $s_l^\sharp = (1-\alpha)-\alpha\frac{h_l}{\delta_l e^+(s_l^\sharp)}$*, and when* $E(e|e<e^+(0)) \equiv \frac{\int_0^{e^+(0)} ef(e)de}{F(e^+(0))} > \max\left\{\frac{\alpha}{1-\alpha}\frac{h_l}{\delta_l}, \frac{\alpha e^+(0)}{1+\alpha}\right\}$*,* $\frac{d(w_n h_n)}{ds_l} > 0$ *for* $s_l \le s_l^\flat$*, where* $s_l^\flat \in (0, s_l^\sharp)$ *satisfies* $s_l^\flat = (1-\alpha) - \alpha\frac{h_l}{\delta_l E(e|e<e^+(s_l^\flat))}$*.*

(ii) (a) If $\overline{e} \le \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$*,* $\frac{d(w_l h_l)}{ds_l} > 0$ *for* $s_l \le s_{l,h}^{\triangle}(e) \in (s_l^\star(e), s_{l,h}^{\triangledown}(e))$ *when* $e > \max\left\{\alpha\left(\frac{h_l}{\delta_l}+e^+(0)\right), \Lambda(\overline{e})\right\}$*,* $\frac{d(w_l h_l)}{ds_l} < 0$ *for* $s_l \ge s_{l,h}^{\triangledown}(e)$ *when* $e > \max\left\{\alpha\left[\frac{h_l}{\delta_l}+E(e|e<e^+(0))\right], \Omega(\overline{e})\right\}$ $(\Omega(\overline{e}) < \Lambda(\overline{e}))$*, and* $\frac{d(w_l h_l)}{ds_l} < 0$ *for any* s_l *when* $e \le \alpha\left[\frac{h_l}{\delta_l}+E(e|e<e^+(0))\right]$*, where* $s_{l,h}^{\triangle}(e)$ *is the greater solution of* $L(s_l) \equiv -ee^+(s_l)s_l^2 + \left[(1-\alpha)e^+(s_l)-(1+\alpha)\frac{h_l}{\delta_l}\right]es_l + \left[-\alpha\left(\frac{h_l}{\delta_l}+e^+(s_l)\right)+e\right]\frac{h_l}{\delta_l} = 0$ *and* $\Lambda(\overline{e}) \equiv \frac{\alpha\left(\frac{h_l}{\delta_l}+\overline{e}\right)}{1+\frac{\delta_l}{4h_l\overline{e}}\left[(1-\alpha)\overline{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]^2}$*, 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, $\frac{d(w_l h_l)}{ds_l} > 0$ *for* $s_l \in \left[\max\{0, s_{l,l}^{\triangle}(e)\}, s_{l,h}^{\triangle}(e)\right]$ *when* $e > \max\left\{\alpha\left(\frac{h_l}{\delta_l}+e^+(0)\right), \Lambda(\overline{e})\right\}$*,*[45] $\frac{d(w_l h_l)}{ds_l} < 0$ *for* $s_l \le \max\{0, s_{l,l}^{\triangledown}(e)\}$ *and* $s_l \ge s_{l,h}^{\triangledown}(e)$ *when* $e \ge \max\left\{\alpha\left[\frac{h_l}{\delta_l}+E(e|e<e^+(0))\right], \Omega(\overline{e})\right\}$*,*[46]

[42] Note that $s_l^\flat$ increases with δ_l, $\underline{s_l}(f)$ decreases with A, δ_n, and δ_l, and $s_{l,h}^{\triangle}(e)$ is defined in Proposition A1 (ii)(a) and does not depend on A, δ_n, and δ_l.

[43] The RHS of the first condition increases with $s_l^\flat$ that increases with δ_l and the expression inside the curly bracket of the second condition is large positive when δ_l is sufficiently large.

[44] 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 the directions of change.

[45] $s_{l,l}^{\triangle}(e) < 0$ when $e \ge \alpha\left(\frac{h_l}{\delta_l}+\overline{e}\right)$.

[46] $s_{l,l}^{\triangledown}(e) < 0$ when $e \ge \alpha\left[\frac{h_l}{\delta_l}+E(e|e<\overline{e})\right]$ and when $E(e|e<\overline{e}) \le \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$. $s_{l,l}^{\triangledown}(e) < s_{l,l}^{\triangle}(e) < s_{l,h}^{\triangle}(e) < s_{l,h}^{\triangledown}(e)$ and $\Omega(\overline{e}) < \Lambda(\overline{e})$ hold.

and $\frac{d(w_l h_l)}{ds_l} < 0$ *for any* s_l *when* $e \leq \Omega(e^+(0))$ *and when* $E(e|e<e^+(0)) \leq \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$ *and* $e \leq \alpha\left[\frac{h_l}{\delta_l}+E(e|e<e^+(0))\right]$, *where* $s^{\triangle}_{l,l}(e)$ $(s^{\nabla}_{l,l}(e))$ *is the smaller 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 C: Proofs of Lemmas and Propositions (Possibly online publication)

Proof of Proposition 1. (i)(a) Suppose that education is worthwhile for local jobs. Consider case $e^+ < \overline{e}$ first. From (2)−(4) and (7), the marginal return to education for local jobs when $e^+ < \overline{e}$ equals

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

In the above equation,

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

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

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

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

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

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

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

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

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

$$=\frac{(s_l)^{\frac{1}{\alpha}}\delta_n e^+}{[h_l(e^+,s_l)]^2}\frac{\begin{pmatrix}h_l(e^+,s_l)\left\{\frac{1-s_l-\alpha s_l}{\alpha s_l}\frac{(1-\alpha)\underline{h}_l}{e^+}\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]+\left(\frac{1-s_l-\alpha s_l}{\alpha s_l}\underline{h}_l+\frac{1-s_l-\alpha}{\alpha}\delta_l e^+\right)e^+f(e^+)\right\}\\-\delta_l(1-s_l)\underline{h}_l\alpha\int_0^{e^+}ef(e)de\end{pmatrix}}{(1-\alpha)\frac{\underline{h}_l}{e^+}\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]+h_l(e^+,s_l)e^+f(e^+)} \quad (17)$$

Since $(1-\alpha)\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\,\overline{e}\right]h_l(e^+,s_l)=\alpha\left[\int_0^{e^+}(\underline{h}_l+\delta_l s_l e)f(e)de\right]e^+$ from (7), (5), (3), and (4), the numerator of the above equation becomes $(s_l)^{\frac{1}{\alpha}}\delta_n e^+$ times

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

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 $s_l\le 1-\alpha$ and negative for $s_l\ge\frac{1}{1+\alpha}$. (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 $(1-s_l-\alpha s_l)\underline{h}_l+(1-s_l-\alpha)s_l\delta_l e^+=0$ too, because, for such s_l, $(1-s_l-\alpha s_l)\underline{h}_l+(1-s_l-\alpha)s_l\delta_l e^+\ge 0$ and $(1-s_l-\alpha s_l)\underline{h}_lF(e^+)+(1-s_l-\alpha)\,\delta_l s_l\int_0^{e^+}ef(e)de\ge-(1-s_l-\alpha)\delta_l s_l\left[e^+F(e^+)-\int_0^{e^+}ef(e)de\right]>0$, where the former statement is true from

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

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

$$\underline{h}_l \frac{d\left[(1-s_l-\alpha s_l)\underline{h}_l F(e^+)+(1-s_l-\alpha)\delta_l s_l \int_0^{e^+} ef(e)de\right]}{ds_l} + \frac{1}{\alpha}\frac{d\left[(1-s_l-\alpha s_l)\underline{h}_l+(1-s_l-\alpha)s_l\delta_l e^+\right]}{ds_l} h_l(e^+,s_l)e^+ f(e^+)$$
$$+\frac{1}{\alpha}\left[(1-s_l-\alpha s_l)\underline{h}_l+(1-s_l-\alpha)s_l\delta_l e^+\right]\frac{d\left[h_l(e^+,s_l)e^+f(e^+)\right]}{ds_l}, \tag{20}$$

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

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

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 $(1-s_l-\alpha s_l)\underline{h}_l+(1-s_l-\alpha)s_l\delta_l e^+ < 0$ 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 $w_l\delta_l s_l - 1$ 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 $w_l\delta_l s_l - 1 > 0$ holds at s_l such that (18) equals zero, there exist two critical values of s_l satisfying $w_l\delta_l s_l - 1 = 0$ 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. Next, consider case $e^+ = \overline{e}$. In this case, from (9), (8), and (10),

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

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

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

In the above equation,

$$\frac{d\left(\frac{\delta_n(1-s_l)\overline{e}}{\underline{h}_l+\delta_l s_l\overline{e}}(s_l)^{\frac{1}{\alpha}}\right)}{ds_l} = \frac{d\left(\frac{\delta_n(1-s_l)\overline{e}}{\underline{h}_l+\delta_l s_l\overline{e}}\right)}{ds_l}(s_l)^{\frac{1}{\alpha}} + \frac{\delta_n(1-s_l)\overline{e}}{\underline{h}_l+\delta_l s_l\overline{e}}\frac{(s_l)^{\frac{1}{\alpha}}}{\alpha s_l} = \frac{(s_l)^{\frac{1}{\alpha}}\delta_n\overline{e}}{\underline{h}_l+\delta_l s_l\overline{e}}\frac{-(\underline{h}_l+\delta_l\overline{e})+(\underline{h}_l+\delta_l s_l\overline{e})\frac{1-s_l}{\alpha s_l}}{\underline{h}_l+\delta_l s_l\overline{e}}$$
$$= \frac{(s_l)^{\frac{1}{\alpha}}\delta_n\overline{e}\frac{1}{\alpha s_l}}{\underline{h}_l+\delta_l s_l\overline{e}}\frac{(1-s_l-\alpha s_l)\underline{h}_l+\delta_l\overline{e}s_l(1-s_l-\alpha)}{\underline{h}_l+\delta_l s_l\overline{e}}, \tag{24}$$

which is positive (negative) for s_l smaller (greater) than the critical value satisfying $(1-s_l-\alpha s_l)\underline{h}_l + \delta_l\overline{e}s_l(1-s_l-\alpha) = 0$. 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 (7), (5), (3), and (4), e^+ is a solution to

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

Thus, e^+ does not depend on A and δ_n and the result on these variables is straightforward from

(14). e^+ depends positively on δ_l from (35) in the proof of Lemma 1 and the derivative of the $RHS - LHS$ of the above equation with respect to δ_l, which equals

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

The result on δ_l is clear from (14) and $\frac{de^+}{d\delta_l} > 0$. (ii) When the return to education for local jobs is positive, $e^+ = \overline{e}$ iff (10) satisfies $\pi_n \leq 1$, i.e.,

$$\begin{aligned}
&\alpha\int_0^{\overline{e}} h_l(e,s_l)f(e)de \leq (1-\alpha)[1-F(\overline{e})]h_l(\overline{e},s_l)\\
\Leftrightarrow\ &\alpha\int_0^{\overline{e}}\left(\underline{h}_l+\delta_l s_l e\right)f(e)de \leq (1-\alpha)[1-F(\overline{e})]\left(\underline{h}_l+\delta_l s_l\overline{e}\right)\\
\Leftrightarrow\ &\underline{h}_l F(\overline{e})+\delta_l s_l\left[\alpha\int_0^{\overline{e}} ef(e)de+(1-\alpha)\overline{e}F(\overline{e})\right] \leq (1-\alpha)\left(\underline{h}_l+\delta_l s_l\overline{e}\right)\\
\Leftrightarrow\ &F(\overline{e}) \leq \frac{(1-\alpha)\left(\underline{h}_l+\delta_l s_l\overline{e}\right)-\delta_l s_l\alpha\int_0^{\overline{e}} ef(e)de}{\underline{h}_l+\delta_l s_l(1-\alpha)\overline{e}} = 1-\alpha\frac{\underline{h}_l+\delta_l s_l\int_0^{\overline{e}} ef(e)de}{\underline{h}_l+\delta_l s_l(1-\alpha)\overline{e}}.
\end{aligned}\tag{27}$$

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

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

$\pi_n \in (0,1]$ 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 $\pi_n = 1$:

$$\begin{aligned}
&\left[\alpha A\left(\frac{F(\overline{e})\underline{h}_l}{[1-F(\overline{e})]h_n(\overline{e},s_l)}\right)^{1-\alpha}\delta_n(1-s_l)-1\right]\overline{e} \leq (1-\alpha)A\left(\frac{[1-F(\overline{e})]h_n(\overline{e},s_l)}{F(\overline{e})\underline{h}_l}\right)^{\alpha}\underline{h}_l\\
\Leftrightarrow\ &\left[\alpha A\left(\frac{F(\overline{e})\underline{h}_l}{[1-F(\overline{e})]\overline{e}}\right)^{1-\alpha}-(\delta_n(1-s_l))^{-\alpha}\right]\overline{e} \leq (1-\alpha)A\left(\frac{[1-F(\overline{e})]\overline{e}}{F(\overline{e})\underline{h}_l}\right)^{\alpha}\underline{h}_l\\
\Leftrightarrow\ &(\delta_n(1-s_l))^{-\alpha} \geq A\left(\frac{[1-F(\overline{e})]\overline{e}}{F(\overline{e})\underline{h}_l}\right)^{\alpha}\frac{\underline{h}_l}{[1-F(\overline{e})]\overline{e}}[F(\overline{e})-(1-\alpha)]\\
\Leftrightarrow\ &\frac{F(\overline{e})-(1-\alpha)}{[F(\overline{e})]^{\alpha}[1-F(\overline{e})]^{1-\alpha}}A(\delta_n(1-s_l))^{\alpha}\left(\frac{\underline{h}_l}{\overline{e}}\right)^{1-\alpha} \leq 1.
\end{aligned}\tag{29}$$

Clearly, the condition is satisfied 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

$$\begin{aligned}
&\frac{[F(\overline{e})]^{\alpha}[1-F(\overline{e})]^{1-\alpha}f(\overline{e})\left\{1-[F(\overline{e})-(1-\alpha)]\left(\frac{\alpha}{F(\overline{e})}-\frac{1-\alpha}{1-F(\overline{e})}\right)\right\}}{\left\{[F(\overline{e})]^{\alpha}[1-F(\overline{e})]^{1-\alpha}\right\}^2} = \frac{[F(\overline{e})]^{\alpha}[1-F(\overline{e})]^{1-\alpha}f(\overline{e})\left\{1-\frac{[F(\overline{e})-(1-\alpha)][\alpha-F(\overline{e})]}{F(\overline{e})[1-F(\overline{e})]}\right\}}{\left\{[F(\overline{e})]^{\alpha}[1-F(\overline{e})]^{1-\alpha}\right\}^2}\\
&= \frac{[F(\overline{e})]^{\alpha}[1-F(\overline{e})]^{1-\alpha}f(\overline{e})\alpha(1-\alpha)}{\left\{[F(\overline{e})]^{\alpha}[1-F(\overline{e})]^{1-\alpha}\right\}^3} > 0.
\end{aligned}\tag{30}$$

□

Proof of Lemma 1. **[When the return to educational investment for local jobs is positive]** When $e^+ < \overline{e}$, by totally differentiating (7), one obtains

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

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

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

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

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

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

$$\begin{aligned}&\alpha H_l(e^+,s_l)\frac{\partial h_n(e^+,s_l)}{\partial e^+}-(1-\alpha)H_n(e^+,s_l)\frac{\partial h_l(e^+,s_l)}{\partial e^+}+h_n(e^+,s_l)h_l(e^+,s_l)f(e^+)\\&=\frac{1}{e^+}\left\{\alpha H_l(e^+,s_l)h_n(e^+,s_l)-(1-\alpha)H_n(e^+,s_l)\left[h_l(e^+,s_l)-\underline{h}_l\right]\right\}+h_n(e^+,s_l)h_l(e^+,s_l)f(e^+)\\&=\frac{1}{e^+}(1-\alpha)H_n(e^+,s_l)\underline{h}_l+h_n(e^+,s_l)h_l(e^+,s_l)f(e^+)\ \text{(from (7))}\\&=\delta_n(1-s_l)\left\{(1-\alpha)\frac{\underline{h}_l}{e^+}\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]+e^+h_l(e^+,s_l)f(e^+)\right\}>0, \quad (35)\end{aligned}$$

The terms of ds_l equals ($\delta\equiv\frac{\delta_n}{\delta_l}$)

$$\begin{aligned}&\alpha H_l(e^+,s_l)\frac{\partial h_n(e^+,s_l)}{\partial s_l}-(1-\alpha)H_n(e^+,s_l)\frac{\partial h_l(e^+,s_l)}{\partial s_l}\\&+\alpha h_n(e^+,s_l)\int_0^{e^+}\frac{\partial h_l(e,s_l)}{\partial s_l}f(e)de-(1-\alpha)\left[\int_{e^+}^{\overline{e}}\frac{\partial h_n(e,s_l)}{\partial s_l}f(e)de+[1-F(\overline{e})]\frac{\partial h_n(\overline{e},s_l)}{\partial s_l}\right]h_l(e^+,s_l)\\&=\delta_l\left\{-e^+\left[\alpha H_l(e^+,s_l)\delta+(1-\alpha)H_n(e^+,s_l)\right]+\alpha h_n(e^+,s_l)\int_0^{e^+}ef(e)de+(1-\alpha)\delta\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]h_l(e^+,s_l)\right\}\\&=\delta_l\left\{\begin{array}{c}-\frac{e^+(1-\alpha)H_n(e^+,s_l)}{h_n(e^+,s_l)}\left[\delta h_l(e^+,s_l)+h_n(e^+,s_l)\right]\\+\alpha h_s(e^+,s_l)\int_0^{e^+}ef(e)de+(1-\alpha)\delta\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]h_l(e^+,s_l)\end{array}\right\}\ \text{(from (7))}\\&=\delta_l\left\{\begin{array}{c}-(1-\alpha)\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]\left[\delta h_l(e^+,s_l)+h_n(e^+,s_l)\right]\\+\alpha h_s(e^+,s_l)\int_0^{e^+}ef(e)de+(1-\alpha)\delta\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]h_l(e^+,s_l)\end{array}\right\}\ \text{(from (5) and (4))}\\&=-\delta_lh_n(e^+,s_l)\left\{(1-\alpha)\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]-\alpha\int_0^{e^+}ef(e)de\right\}<0, \quad (36)\end{aligned}$$

where the last inequality holds because

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

$$\Leftrightarrow\left\{\alpha F(e^+)e^+-(1-\alpha)\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]\right\}\underline{h}_l=s_l\delta_le^+\left\{(1-\alpha)\left[\int_{e^+}^{\overline{e}}ef(e)de+[1-F(\overline{e})]\overline{e}\right]-\alpha\int_0^{e^+}ef(e)de\right\}\ \text{from (7)}$$

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,

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

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

When $e^+ = \overline{e}$, from (10),

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

[**When the return is negative**] First consider case $e^+ < \overline{e}$. From (12), (4), and (11), $w_n h_n(e^+,s_l) - e^+ = w_l\underline{h}_l$ can be expressed as

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

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

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

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

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

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,

$$\frac{\partial w_n}{\partial e^+} = (1-\alpha)w_n\left[\frac{1}{H_l(e^+,s_l)}\frac{\partial H_l(e^+,s_l)}{\partial e^+} - \frac{1}{H_n(e^+,s_l)}\frac{\partial H_n(e^+,s_l)}{\partial e^+}\right]$$

$$= (1-\alpha)w_n f(e^+)\left[\frac{h_l(e^+,s_l)}{H_l(e^+,s_l)} + \frac{h_n(e^+,s_l)}{H_n(e^+,s_l)}\right]$$

$$= w_n f(e^+)\frac{h_n(e^+,s_l)}{H_n(e^+,s_l)} \text{ (from (7))}$$

$$= \frac{w_n e^+}{\int_{e^+}^{\overline{e}} ef(e)de + [1-F(\overline{e})]\overline{e}}f(e^+) > 0. \tag{41}$$

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

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

The last equality is because $\alpha\left[\underline{h}_l\int_0^{e^+} f(e)de + \delta_l s_l\int_0^{e^+} ef(e)de\right]e^+ = (1-\alpha)\left[\int_{e^+}^{\overline{e}} ef(e)de + [1-F(\overline{e})]\overline{e}\right]h_l(e^+,s_l)$ from (7). From (41), (42), and (37) in the proof of Lemma 1,

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

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

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

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

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

where the last equality is again from (7). When $e^+ = \overline{e}$, from (8), (9), and (10),

$$
\begin{aligned}
H_l(\pi_n, s_l) &= \int_0^{\overline{e}} h_l(e,s_l)f(e)de + \left(1-\alpha\left\{1 + \tfrac{\int_0^{\overline{e}} h_l(e,s_l)f(e)de}{[1-F(\overline{e})]h_l(\overline{e},s_l)}\right\}\right)[1-F(\overline{e})]\,h_l(\overline{e},s_l) \\
&= (1-\alpha)\left\{\int_0^{\overline{e}} h_l(e,s_l)f(e)de + [1-F(\overline{e})]\,h_l(\overline{e},s_l)\right\}, \qquad (44) \\
H_n(\pi_n, s_l) &= \alpha\tfrac{h_n(\overline{e},s_l)}{h_l(\overline{e},s_l)}\left\{\int_0^{\overline{e}} h_l(e,s_l)f(e)de + [1-F(\overline{e})]\,h_l(\overline{e},s_l)\right\}. \qquad (45)
\end{aligned}
$$

By substituting the above equations into (2),

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

Thus,

$$\frac{dw_n}{ds_l} = (1-\alpha)w_n\left(\tfrac{1}{h_l(\bar{e},s_l)}\delta_l\bar{e} + \tfrac{1}{h_n(\bar{e},s_l)}\delta_n\bar{e}\right)$$
$$= (1-\alpha)w_n\tfrac{1}{1-s_l}\tfrac{h_l(\bar{e},s_l)+\frac{\delta_l}{\delta_n}h_n(\bar{e},s_l)}{h_l(\bar{e},s_l)} > 0. \quad (47)$$

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 $e^+ < \bar{e}$ and of (40) when $e^+ = \bar{e}$. □

Proof of Proposition 2. When $e^+ < \bar{e}$, $w_l\underline{h}_l = w_n h_n(e^+, s_l) - e^+ = [w_n h_n(1, s_l) - 1]\, e^+$ 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 $e > e^+$ also 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,

$$\frac{d[w_n h_n(\bar{e}, s_l)]}{ds_l} = \frac{dw_n}{ds_l}h_n(\bar{e}, s_l) + w_n\frac{dh_n(\bar{e}, s_l)}{ds_l}$$
$$= \tfrac{w_n h_n(\bar{e},s_l)}{1-s_l}\left[(1-\alpha)\tfrac{h_l(\bar{e},s_l)+\frac{\delta_l}{\delta_n}h_n(\bar{e},s_l)}{h_l(\bar{e},s_l)} - 1\right]$$
$$= \tfrac{w_n h_n(\bar{e},s_l)}{1-s_l}\tfrac{(1-\alpha)\frac{\delta_l}{\delta_n}h_n(\bar{e},s_l)-\alpha h_l(\bar{e},s_l)}{h_l(\bar{e},s_l)}. \quad (48)$$

Thus,

$$\frac{d(w_n h_n)}{ds_l} \gtreqless 0 \Leftrightarrow (1-\alpha)\delta_l(1-s_l)\bar{e} - \alpha\left(\underline{h}_l + \delta_l s_l \bar{e}\right) \gtreqless 0$$
$$\Leftrightarrow s_l \lesseqgtr (1-\alpha) - \alpha\tfrac{\underline{h}_l}{\delta_l \bar{e}}. \quad (49)$$

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

$$\frac{d[w_l h_l(e, s_l)]}{ds_l} = \frac{dw_l}{ds_l}h_l(e, s_l) + w_l\frac{dh_l(e, s_l)}{ds_l}$$
$$= -\tfrac{\alpha}{1-\alpha}\tfrac{w_l}{w_n}\tfrac{dw_n}{ds_l}h_l(e, s_l) + w_l\tfrac{dh_l(e,s_l)}{ds_l}$$
$$= \tfrac{w_l}{1-s_l}\tfrac{1}{h_l(\bar{e},s_l)}\left\{-\alpha\left[h_l(\bar{e}, s_l) + \tfrac{\delta_l}{\delta_n}h_n(\bar{e}, s_l)\right]h_l(e, s_l) + \delta_l e(1-s_l)h_l(\bar{e}, s_l)\right\}. \quad (50)$$

Thus,

$$\frac{d(w_l h_l)}{ds_l} \gtreqless 0 \Leftrightarrow -\alpha\left(\underline{h}_l + \delta_l\bar{e}\right)\left(\underline{h}_l + \delta_l s_l e\right) + \delta_l e(1-s_l)\left(\underline{h}_l + \delta_l s_l \bar{e}\right) \gtreqless 0$$
$$\Leftrightarrow -e\bar{e}(s_l)^2 + \left[(1-\alpha)\bar{e} - (1+\alpha)\tfrac{\underline{h}_l}{\delta_l}\right]es_l + \left[-\alpha\left(\tfrac{\underline{h}_l}{\delta_l} + \bar{e}\right) + e\right]\tfrac{\underline{h}_l}{\delta_l} \gtreqless 0. \quad (51)$$

(a) Suppose that the LHS of (51) is positive at $s_l = 0$, i.e., $e > \alpha\left(\frac{\underline{h}_l}{\delta_l} + \bar{e}\right)$. Because the derivative of the LHS at $s_l = 0$ is non-positive, i.e., $(1-\alpha)\bar{e} - (1+\alpha)\frac{\underline{h}_l}{\delta_l} \leq 0$ and the LHS at $s_l = 1$ equals

$-e\bar{e}+\left[(1-\alpha)\bar{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]e+\left[-\alpha\left(\frac{h_l}{\delta_l}+\bar{e}\right)+e\right]\frac{h_l}{\delta_l}=-\alpha\left(\frac{h_l}{\delta_l}+\bar{e}\right)\left(\frac{h_l}{\delta_l}+e\right)<0$, there exists an $s_l^{\star}(e)\in(0,1)$ such that $\frac{d(w_lh_l)}{ds_l}\gtreqless 0\Leftrightarrow s_l\lesseqgtr s_l^{\star}(e)$, where $s_l^{\star\prime}(e)>0$, since, from (51),

$$s_l^{\star}(e)=\frac{\left[(1-\alpha)\bar{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]e+\sqrt{\left[(1-\alpha)\bar{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]^2e^2+4e\bar{e}\left[-\alpha\left(\frac{h_l}{\delta_l}+\bar{e}\right)+e\right]\frac{h_l}{\delta_l}}}{2e\bar{e}}$$
$$=\frac{\left[(1-\alpha)\bar{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]+\sqrt{\left[(1-\alpha)\bar{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]^2+4\bar{e}\left[-\alpha\left(\frac{h_l}{\delta_l}+\bar{e}\right)e^{-1}+1\right]\frac{h_l}{\delta_l}}}{2\bar{e}}. \tag{52}$$

$s_l^{\star}(e)<s_l^{\star\star}$ for $e<\bar{e}$, since $w_lh_l(\bar{e},s_l)=w_nh_n(\bar{e},s_l)$ (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., $e\leq\alpha\left(\frac{h_l}{\delta_l}+\bar{e}\right)$. Because the derivative of the LHS at $s_l=0$ is non-positive and the LHS at $s_l=1$ is negative, $\frac{d(w_lh_l)}{ds_l}<0$ for any $s_l>0$ (and $\frac{d(w_lh_l)}{ds_l}<(=)0$ at $s_l=0$ when $e<(=)\alpha\left(\frac{h_l}{\delta_l}+\bar{e}\right)$). (b) The case in which the LHS of (51) at $s_l=0$ is positive, i.e., $e>\alpha\left(\frac{h_l}{\delta_l}+\bar{e}\right)$, can be proven as in (a). Suppose that the LHS at $s_l=0$ is zero, i.e., $e=\alpha\left(\frac{h_l}{\delta_l}+\bar{e}\right)$. Since the derivative of the LHS at $s_l=0$ is positive, i.e., $(1-\alpha)\delta_l\bar{e}-(1+\alpha)\underline{h}_l>0$, 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, $\frac{d(w_lh_l)}{ds_l}\gtreqless 0$ for positive $s_l\lesseqgtr s_l^{\star}(e)$ and $\frac{d(w_lh_l)}{ds_l}=0$ at $s_l=0$. Instead, suppose that the LHS of (51) at $s_l=0$ is negative, $e<\alpha\left(\frac{h_l}{\delta_l}+\bar{e}\right)$. Since the derivative of the LHS at $s_l=0$ is positive, i.e., $(1-\alpha)\delta_l\bar{e}-(1+\alpha)\underline{h}_l>0$, the LHS is positive (negative) when $e>(<)\frac{\alpha\left(\frac{h_l}{\delta_l}+\bar{e}\right)\frac{h_l}{\delta_l}}{\left\{-\bar{e}s_l+\left[(1-\alpha)\bar{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]\right\}s_l+\frac{h_l}{\delta_l}}$, where $\frac{\alpha\left(\frac{h_l}{\delta_l}+\bar{e}\right)\frac{h_l}{\delta_l}}{\left\{-\bar{e}s_l+\left[(1-\alpha)\bar{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]\right\}s_l+\frac{h_l}{\delta_l}}<\alpha\left(\frac{h_l}{\delta_l}+\bar{e}\right)$ when $s_l>0$ and $-\bar{e}s_l+\left[(1-\alpha)\bar{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]>0\Leftrightarrow s_l\in(0,(1-\alpha)-(1+\alpha)\frac{\frac{h_l}{\delta_l}}{\bar{e}})$. So the LHS of (51) is negative for any $e<\alpha\left(\frac{h_l}{\delta_l}+\bar{e}\right)$ when $s_l\geq(1-\alpha)-(1+\alpha)\frac{\frac{h_l}{\delta_l}}{\bar{e}}$. For $s_l<(1-\alpha)-(1+\alpha)\frac{\frac{h_l}{\delta_l}}{\bar{e}}$, the LHS of (51) is positive (negative) when $e>(<)\frac{\alpha\left(\frac{h_l}{\delta_l}+\bar{e}\right)\frac{h_l}{\delta_l}}{\left\{-\bar{e}s_l+\left[(1-\alpha)\bar{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]\right\}s_l+\frac{h_l}{\delta_l}}$, where the RHS is lowest when $-\bar{e}s_l+\left[(1-\alpha)\bar{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]-\bar{e}s_l=0\Leftrightarrow s_l=\frac{(1-\alpha)\bar{e}-(1+\alpha)\frac{h_l}{\delta_l}}{2\bar{e}}$. Hence, the LHS of (51) is negative, i.e., $\frac{d(w_lh_l)}{ds_l}<0$, for any s_l when $e\leq\left[\frac{\alpha\left(\frac{h_l}{\delta_l}+\bar{e}\right)\frac{h_l}{\delta_l}}{\left\{-\bar{e}s_l+\left[(1-\alpha)\bar{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]\right\}s_l+\frac{h_l}{\delta_l}}\text{ at } s_l=\frac{(1-\alpha)\bar{e}-(1+\alpha)\frac{h_l}{\delta_l}}{2\bar{e}}\right]=\frac{\alpha\left(\frac{h_l}{\delta_l}+\bar{e}\right)\frac{h_l}{\delta_l}}{\frac{h_l}{\delta_l}+\frac{\left[(1-\alpha)\bar{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]^2}{4\bar{e}}}=\frac{\alpha\left(\frac{h_l}{\delta_l}+\bar{e}\right)}{1+\frac{\delta_l}{4\underline{h}\bar{e}}\left[(1-\alpha)\bar{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]^2}$, except at $s_l=\frac{(1-\alpha)\bar{e}-(1+\alpha)\frac{h_l}{\delta_l}}{2\bar{e}}$ and $e=\frac{\alpha\left(\frac{h_l}{\delta_l}+\bar{e}\right)}{1+\frac{\delta_l}{4\underline{h}\bar{e}}\left[(1-\alpha)\bar{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]^2}$, in which the derivative is zero. When $e\in\left(\frac{\alpha\left(\frac{h_l}{\delta_l}+\bar{e}\right)}{1+\frac{\delta_l}{4\underline{h}\bar{e}}\left[(1-\alpha)\bar{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]^2},\alpha\left(\frac{h_l}{\delta_l}+\bar{e}\right)\right)$, there exist $s_l^{\circ}(e)$ and $s_l^{\star}(e)$, where $0<s_l^{\circ}(e)<s_l^{\star}(e)<(1-\alpha)-(1+\alpha)\frac{\frac{h_l}{\delta_l}}{\bar{e}}<s_l^{\star\star}$, such that the LHS of (51) equals 0, and the LHS is negative for $s_l<s_l^{\circ}(e)$ and $s_l>s_l^{\star}(e)$ and positive for $s_l\in(s_l^{\circ}(e),s_l^{\star}(e))$. $s_l^{\star\prime}(e)>0>s_l^{\circ\prime}(e)$ from (52). □

Proof of Proposition 4. $e^+ = \bar{e}$ always holds when $F(\bar{e}) \le 1-\alpha$ from Proposition 1 (ii). The following lemma is used in the proof of the proposition.

Lemma A2. *When $e^+ = \bar{e}$, 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 $e^+ = \bar{e}$ and the return is negative, $w_n h_n(\bar{e}, s_l) - \bar{e} = w_l h_l(0, s_l) \Leftrightarrow \left[\alpha A\left(\frac{H_l}{H_n}\right)^{1-\alpha}\delta_n(1-s_l)-1\right]\bar{e} = (1-\alpha)A\left(\frac{H_n}{H_l}\right)^{\alpha}\underline{h}_l$, where $H_l = \{F(\bar{e})+(1-\pi_n)[1-F(\bar{e})]\}\,\underline{h}_l$ and $H_n = \pi_n[1-F(\bar{e})]h_n(\bar{e}, s_l)$, from (13) and (12). When $e^+ = \bar{e}$ and the return is positive, $w_n h_n(\bar{e}, s_l) - \bar{e} = w_l h_l(\bar{e}, s_l) - \bar{e} \Leftrightarrow \left[\alpha A\left(\frac{H_l}{H_n}\right)^{1-\alpha}\delta_n(1-s_l)-1\right]\bar{e} = (1-\alpha)A\left(\frac{H_n}{H_l}\right)^{\alpha}(\underline{h}_l+\delta_l s_l \bar{e}) - \bar{e}$, where $H_l = \int_0^{\bar{e}} h_l(e, s_l)f(e)de + (1-\pi_n)[1-F(\bar{e})]h_l(\bar{e}, s_l)$ and $H_n = \pi_n[1-F(\bar{e})]h_n(\bar{e}, s_l)$, from (2), (8), and (9). When the return is zero, i.e., $w_l\delta_l s_l - 1 = 0$, this equation becomes $\left[\alpha A\left(\frac{H_l}{H_n}\right)^{1-\alpha}\delta_n(1-s_l)-1\right]\bar{e} = (1-\alpha)A\left(\frac{H_n}{H_l}\right)^{\alpha}\underline{h}_l$, 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, net earnings when the return is zero are same as net earnings when the return is negative and $s_l \to \underline{s_l}, \overline{s_l}$. $\square$

(i) Earnings decrease with s_l when the return to education for local jobs is negative from Proposition 2, while when it is positive and $e^+ = \bar{e}$, earnings of all decrease with s_l for $s_l > s_l^{\star\star} \equiv (1-\alpha) - \alpha\frac{\underline{h}_l}{\delta_l \bar{e}}$ from Proposition 3. From Proposition 1 (i)(b), the lower critical value for the negative return, $\underline{s_l}$, decreases with A, δ_l, and δ_n. Hence, from Lemma A2, net earnings of all decrease with s_l when A, δ_l, and δ_n are small enough that $\underline{s_l} \ge s_l^{\star\star} = s_l^\star(e^+) = s_l^\star(\bar{e}) \equiv (1-\alpha) - \alpha\frac{\underline{h}_l}{\delta_l \bar{e}}$. Note that $s_l^{\star\star}$ is smaller than the higher critical value $\overline{s_l}$, since $s_l^{\star\star} < 1-\alpha < \overline{s_l}$ 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 $s_l > s_l^{\star\star}$. Thus, 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

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

where π_n is a solution to (28) in the proof of Proposition 1. From (2) and (22) in the proof of Proposition 1, net earnings of such workers at $s_l = s_l^{\star\star}$ equal

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

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

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

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

$$\frac{\alpha}{\frac{h_l}{\delta_l \overline{e}}} - \frac{\left(\frac{h_l}{\delta_l \overline{e}}\right)^{-2}}{1+\left(\frac{h_l}{\delta_l \overline{e}}\right)^{-1}} = \frac{\alpha-(1-\alpha)\left(\frac{h_l}{\delta_l \overline{e}}\right)^{-1}}{\frac{h_l}{\delta_l \overline{e}}\left[1+\left(\frac{h_l}{\delta_l \overline{e}}\right)^{-1}\right]} < 0 \text{ from } \frac{h_l}{\delta_l \overline{e}} < \frac{1-\alpha}{\alpha}. \tag{56}$$

(b) Consider case $\overline{e} \leq \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$ first. Since $s_l^\star(e)$ increases with e, $s_l^\star(\alpha(\frac{h_l}{\delta_l}+\overline{e})) = 0$, and $s_l^\star(\overline{e}) = s_l^{\star\star}$ from (52) in the proof of Proposition 3, there exists an $e^\dagger \in (\alpha(\frac{h_l}{\delta_l}+\overline{e}), \overline{e})$, 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 $\overline{e} > \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$. If A, δ_l, and δ_n are small enough that $s_l^\star(\alpha(\frac{h_l}{\delta_l}+\overline{e})) = 1-\alpha-\frac{(1+\alpha)\frac{h_l}{\delta_l}}{\overline{e}} < \underline{s_l}$, the result is same as the case of $\overline{e} \leq \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$. Let $\Lambda(\overline{e}) \equiv \frac{\alpha\left(\frac{h_l}{\delta_l}+\overline{e}\right)}{1+\frac{\delta_l}{4h\overline{e}}\left[(1-\alpha)\overline{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]^2}$. If $s_l^\star(\alpha(\frac{h_l}{\delta_l}+\overline{e})) > \underline{s_l} > s_l^\star(\Lambda(\overline{e})) = \frac{(1-\alpha)-\frac{(1+\alpha)\frac{h_l}{\delta_l}}{\overline{e}}}{2}$, there exists an $e^\ddagger \in \left(\Lambda(\overline{e}), \alpha\left(\frac{h_l}{\delta_l}+\overline{e}\right)\right)$ such that $s_l^\star(e^\ddagger) = \underline{s_l}$. The results for those with $e > \alpha\left(\frac{h_l}{\delta_l}+\overline{e}\right)$ and those with $e \leq \Lambda(\overline{e})$ are same as the corresponding cases ($e > e^\dagger$ and $e \leq e^\dagger$ respectively) of $\overline{e} \leq \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$. As for those with $e \in \left(\Lambda(\overline{e}), \alpha\left(\frac{h_l}{\delta_l}+\overline{e}\right)\right)$, $s_l^\circ(e) \leq \frac{(1-\alpha)-\frac{(1+\alpha)\frac{h_l}{\delta_l}}{\overline{e}}}{2} = s_l^\star(\Lambda(\overline{e})) < s_l^\star(e^\ddagger) = \underline{s_l}$ holds for any e from (51) and (52) in the proof of Proposition 3. Hence, the result of those with $e > e^\ddagger$ is similar to that of those with $e > \alpha\left(\frac{h_l}{\delta_l}+\overline{e}\right)$, and the result of those with $e \leq e^\ddagger$ is same as that of those with $e \leq \Lambda(\overline{e})$. In sum, the result is similar to when $\overline{e} \leq \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$ except that the critical wealth level is $e^\ddagger$, not $e^\dagger$. Finally, if $s_l^\star(\Lambda(\overline{e})) > \underline{s_l}$, since $s_l^\circ(e)$ decreases with e, $s_l^\circ(\alpha(\frac{h_l}{\delta_l}+\overline{e})) = 0$, and $s_l^\circ(\Lambda(\overline{e})) = s_l^\star(\Lambda(\overline{e})) = \frac{(1-\alpha)-\frac{(1+\alpha)\frac{h_l}{\delta_l}}{\overline{e}}}{2}$ from (51) and (52), there exists an $e^\divideontimes \in \left(\Lambda(\overline{e}), \alpha\left(\frac{h_l}{\delta_l}+\overline{e}\right)\right)$ such that $s_l^\circ(e^\divideontimes) = \underline{s_l}$. Hence, as for workers with $e \in \left(\Lambda(\overline{e}), \alpha\left(\frac{h_l}{\delta_l}+\overline{e}\right)\right)$, net earnings of those with $e < e^\divideontimes$ 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 those with $e \geq e^\divideontimes$ 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 those with $e > \alpha\left(\frac{h_l}{\delta_l}+\overline{e}\right)$ and those with $e \leq \Lambda(\overline{e})$ are same as the previous case. In sum, net earnings of workers with wealth 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. 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

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

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

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

which holds when A, δ_n, and δ_l are large, because π_n increases with A and δ_n, 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 is positive when A and δ_l are large from (55):

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

(c) From the proof of (b), the threshold wealth when $\overline{e} \le \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$ is $e^\dagger \in (\alpha(\frac{h_l}{\delta_l}+\overline{e}), \overline{e})$ such that $s_l^\star(e^\dagger) = \underline{s_l}$. The threshold when $\overline{e} > \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$ is $e^\dagger \in (\alpha(\frac{h_l}{\delta_l}+\overline{e}), \overline{e})$ if $s_l^\star(\alpha(\frac{h_l}{\delta_l}+\overline{e})) < \underline{s_l}$, $e^\ddagger \in \left(\Lambda(\overline{e}), \alpha\left(\frac{h_l}{\delta_l}+\overline{e}\right)\right)$ such that $s_l^\star(e^\ddagger) = \underline{s_l}$ if $s_l^\star(\alpha(\frac{h_l}{\delta_l}+\overline{e})) > \underline{s_l} > s_l^\star(\Lambda(\overline{e}))$, and $\Lambda(\overline{e})$ if $s_l^\star(\Lambda(\overline{e})) > \underline{s_l}$. When $\overline{e} > \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$, case $s_l^\star(\alpha(\frac{h_l}{\delta_l}+\overline{e})) < \underline{s_l}$ is realized when A, δ_l, and δ_n are small, and case $s_l^\star(\Lambda(\overline{e})) > \underline{s_l}$ is realized when they are large, because $\underline{s_l}$ decreases with A, δ_l, and δ_n from Proposition 1, and $s_l^\star(\alpha(\frac{h_l}{\delta_l}+\overline{e})) = (1-\alpha)-(1+\alpha)\frac{h_l}{\delta_l\overline{e}}$ and $s_l^\star(\Lambda(\overline{e})) = \frac{1}{2}\left[(1-\alpha)-(1+\alpha)\frac{h_l}{\delta_l\overline{e}}\right]$ increase with δ_l. Further, $e^\dagger$ and $e^\ddagger$ decrease with A, δ_l, and δ_n, because $s_l^\star(e)$ increases with δ_l at $e = e^\dagger$, $e^\ddagger$ from Lemma A3 below and increases with e, and $\underline{s_l}$ decreases with A, δ_l, and δ_n. Hence, the threshold of wealth decreases with A and δ_n (except when the threshold is $\alpha\left(\frac{h_l}{\delta_l}+\overline{e}\right)$ and $\Lambda(\overline{e})$) and δ_l. □

Lemma A3. $\frac{\partial s_l^\star(e^\dagger)}{\partial\delta_l} > 0$ *and* $\frac{\partial s_l^\star(e^\ddagger)}{\partial\delta_l} > 0$

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

$$\begin{aligned}
&(1+\alpha)\underline{h}_l-\frac{1}{2}\left\{-(1+\alpha)\underline{h}_l 2\left[(1-\alpha)\overline{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]+4\overline{e}\left[-\alpha\left(\frac{h_l}{\delta_l}+\overline{e}\right)(e)^{-1}+1\right]\underline{h}_l-4\overline{e}\alpha\frac{h_l}{\delta_l}(e)^{-1}\underline{h}_l\right\}\\
&\qquad\times\left\{\left[(1-\alpha)\overline{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]^2+4\frac{\overline{e}}{\delta_l}\left[-\alpha\left(\frac{h_l}{\delta_l}+\overline{e}\right)(e)^{-1}+1\right]\underline{h}_l\right\}^{-1/2}.
\end{aligned} \tag{60}$$

Thus,

$$\begin{aligned}
\frac{\partial s_l^\star(e)}{\partial\delta_l}\gtreqless 0 \Leftrightarrow &-(1+\alpha)\underline{h}_l 2\left[(1-\alpha)\overline{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]+4\overline{e}\left[-\alpha\left(\frac{h_l}{\delta_l}+\overline{e}\right)(e)^{-1}+1\right]\underline{h}_l-4\overline{e}\alpha\frac{h_l}{\delta_l}(e)^{-1}\underline{h}_l \lesseqgtr 2(1+\alpha)\underline{h}_l\\
&\times\left\{\left[(1-\alpha)\overline{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]^2+4\frac{\overline{e}}{\delta_l}\left[-\alpha\left(\frac{h_l}{\delta_l}+\overline{e}\right)(e)^{-1}+1\right]\underline{h}_l\right\}^{1/2}.
\end{aligned} \tag{61}$$

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

$$\frac{\partial s_l^\star(e)}{\partial \delta_l} \gtreqless 0 \Leftrightarrow \left\{-(1+\alpha)\underline{h}_l 2\left[(1-\alpha)\overline{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]+4\overline{e}\left[-\alpha\left(\frac{h_l}{\delta_l}+\overline{e}\right)(e)^{-1}+1\right]\underline{h}_l-4\overline{e}\alpha\frac{h_l}{\delta_l}(e)^{-1}\underline{h}_l\right\}^2$$
$$\gtreqless [2(1+\alpha)\underline{h}_l]^2\left\{\left[(1-\alpha)\overline{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]^2+4\frac{\overline{e}}{\delta_l}\left[-\alpha\left(\frac{h_l}{\delta_l}+\overline{e}\right)(e)^{-1}+1\right]\underline{h}_l\right\} \tag{62}$$
$$\Leftrightarrow -(1+\alpha)\left[(1-\alpha)\overline{e}-(1+\alpha)\frac{h_l}{\delta_l}\right]\left[-\alpha\left(2\frac{h_l}{\delta_l}+\overline{e}\right)(e)^{-1}+1\right]+\left[-\alpha\left(2\frac{h_l}{\delta_l}+\overline{e}\right)(e)^{-1}+1\right]^2\overline{e}$$
$$\gtreqless (1+\alpha)^2\frac{1}{\delta_l}\left[-\alpha\left(\frac{h_l}{\delta_l}+\overline{e}\right)(e)^{-1}+1\right]\underline{h}_l$$
$$\Leftrightarrow \left[-\alpha\left(2\frac{h_l}{\delta_l}+\overline{e}\right)(e)^{-1}+1\right]\left[-\left(2\frac{h_l}{\delta_l}+\overline{e}\right)(e)^{-1}+\alpha\right]\overline{e}-(1+\alpha)^2\frac{1}{\delta_l}\left(\frac{h_l}{\delta_l}\right)(e)^{-1}\underline{h}_l \gtreqless 0. \tag{63}$$

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

Proof of Corollary 1. Since $F(\overline{e})=0<1-\alpha$, the proof of Proposition 4 can be applied with $a>\overline{e}$ for anyone. The result is straightforward from the proof, since $e=\overline{e}=e^+$ and thus $s_l^\star(e)=s_l^\star(\overline{e})=s_l^{\star\star}$ for those choosing a local job, and $\overline{e}>\alpha\left(\frac{h_l}{\delta_l}+\overline{e}\right)$ by assumption (see footnote 20). □

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, $w_n h_n(e^+, s_l) - e^+ = w_l \underline{h}_l \Leftrightarrow \left[\alpha A\left(\frac{H_l}{H_n}\right)^{1-\alpha}\delta_n(1-s_l)-1\right]e^+ = (1-\alpha)A\left(\frac{H_n}{H_l}\right)^{\alpha}\underline{h}_l$, 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 (12). When $e^+ < \overline{e}$ and the return is positive, $w_n h_n(e^+, s_l) - e^+ = w_l h_l(e^+, s_l) - e^+ \Leftrightarrow \left[\alpha A\left(\frac{H_l}{H_n}\right)^{1-\alpha}\delta_n(1-s_l)-1\right]e^+ = (1-\alpha)A\left(\frac{H_n}{H_l}\right)^{\alpha}(\underline{h}_l+\delta_l s_l e^+) - e^+$, 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., $w_l\delta_l s_l - 1 = 0$, this equation becomes $\left[\alpha A\left(\frac{H_l}{H_n}\right)^{1-\alpha}\delta_n(1-s_l)-1\right]e^+ = (1-\alpha)A\left(\frac{H_n}{H_l}\right)^{\alpha}\underline{h}_l$, the same equation as the case of the negative return. Because $\frac{H_n}{H_l}$ under the negative return is greater than 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 those with a local job are greater and of those 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 those with a national (local) job increase (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 education for local jobs is 0 when $e^+ < \overline{e}$ be $\underline{s_l}(f)$ ($\overline{s_l}(f)$), whose existence is shown in the proof of Proposition 1. As for those who have abundant wealth and choose a national job for any s_l, i.e., $a \geq \min\{e^+(1), \overline{e}\}$, net earnings decrease with s_l except at $s_l = \underline{s_l}(f)$, where they increase discontinuously from Lemma A4, if $e^+(0) \leq \frac{\alpha}{1-\alpha}\frac{h_l}{\delta_l}$ or $\underline{s_l}(f) \geq s_l^\sharp$ satisfying $s_l^\sharp = (1-\alpha) - \alpha\frac{h_l}{\delta_l e^+(s_l^\sharp)}$ from Propositions A1 (i) and 2. The former condition holds when δ_l is small, because $e^+(0)$ does not depend on A, δ_n, and δ_l from (D3) in the proof of Proposition A1 in Appendix D. The latter condition holds when A, δ_n, and δ_l are small, since $\underline{s_l}(f)$ decreases with A, δ_n, and δ_l from Proposition 1 and is greater than $1-\alpha$ when these variables are small from the part of the proof of Proposition 1 (i)(a) just after Lemma A1, while $s_l^\sharp < 1-\alpha$. As for those who choose a local job for any s_l, i.e., $a = e < e^+(0)$, net earnings decrease with s_l except at $s_l = \overline{s_l}(f)$, where net earnings increase discontinuously from Lemma A4, when $E(e|e< e^+(0)) \leq \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$ and $e \leq \alpha\left(\frac{h_l}{\delta_l} + E(e|e<e^+(0))\right)$, when $E(e|e<e^+(0)) > \frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$ and $e \leq \Omega(e^+(0))$, and when $\underline{s_l}(f) \geq s_{l,h}^\bigtriangledown(e)$ for greater e, where $s_{l,h}^\bigtriangledown(e)$ is the greater solution of $M(s_l) = 0$ ($M(s_l)$ equals $L(s_l)$ with $e^+(s_l)$ replaced with $E(e|e<e^+(s_l)) \equiv \frac{\int_0^{e^+(s_l)} ef(e)de}{F(e^+(s_l))}$), from Propositions A1 (ii) and 2. Thus, irrespective of $a = e$, net earnings decrease with s_l except at $s_l = \overline{s_l}(f)$, if $\underline{s_l}(f) \geq s_{l,h}^\bigtriangledown(e)$. The condition holds when A, δ_n, and δ_l are small, because as shown above, $\underline{s_l}(f) > 1-\alpha$ when they are small, while $1-\alpha \geq s_{l,h}^\bigtriangledown(e)$ holds when δ_l is small. The latter statement is true because $M(1-\alpha) = -\alpha\left[\left(\frac{h_l}{\delta_l} + E(e|e<e^+(1-\alpha))\right) - \alpha e\right]\frac{h_l}{\delta_l} < -\alpha\left(\frac{h_l}{\delta_l} - \alpha e\right)\frac{h_l}{\delta_l} \leq 0$ when δ_l is sufficiently small. Finally, as for those who choose a national (local) job at small (large) s_l, if the above conditions hold, 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 net earnings increase discontinuously from Lemma A4. Under the above condition, 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

$$\left[\alpha A\left(\frac{H_l(e^+,s_l)}{H_n(e^+,s_l)}\right)^{1-\alpha}\delta_n - 1\right]e = \left[\alpha A\left(\frac{1-\alpha}{\alpha}\frac{h_l}{\delta_n e^+(0)}\right)^{1-\alpha}\delta_n - 1\right]e, \tag{64}$$

where the equality sign is from (7). From (14) in the proof of Proposition 1 and (2), their net earnings at $s_l = \underline{s_l}(f)$ equal

$$\left\{\alpha A\left[\frac{(1-\alpha)\left(h_l+\delta_l s_l e^+\right)}{\alpha\delta_n(1-s_l)e^+}\right]^{1-\alpha}\delta_n(1-s_l)-1\right\}e=\left\{\alpha A[(1-\alpha)A\delta_l s_l]^{\frac{1-\alpha}{\alpha}}\delta_n(1-s_l)-1\right\}e \tag{65}$$

$$<\left\{\alpha A\left[(1-\alpha)^2 A\delta_l\right]^{\frac{1-\alpha}{\alpha}}\delta_n\alpha-1\right\}e \text{ (from } \underline{s_l}(f)<1-\alpha), \tag{66}$$

where the equality sign is from the fact that the return for local jobs is 0 at $s_l = \underline{s_l}(f)$. From (64) and (66), the net earnings are maximized at $s_l = 0$ if

$$\left(\frac{1-\alpha}{\alpha}\frac{h_l}{\delta_n e^+(0)}\right)^{1-\alpha}>\left[(1-\alpha)^2 A\delta_l\right]^{\frac{1-\alpha}{\alpha}}\alpha, \tag{67}$$

which holds when A, δ_n, and δ_l are small, since $e^+(0)$ does not depend on A, δ_n, and δ_l. As for those who choose a national job at small s_l and a local job at large s_l, the proof for those who choose a local (national) job for any s_l applies if the switch to a local job occurs at $s_l \leq \underline{s_l}(f)$ ($s_l \geq \overline{s_l}(f)$). If 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, since earnings when the return is negative decrease with s_l from Proposition 2. (ii) From the above results, if $a<\min\left\{e^+(0), \alpha\left(\frac{h_l}{\delta_l}+E(e|e<e^+(0))\right)\right\}$ when $E(e|e<e^+(0))\leq\frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$ or if $a<\min\{e^+(0),\Omega(e^+(0))\}$ when $E(e|e<e^+(0))>\frac{1+\alpha}{1-\alpha}\frac{h_l}{\delta_l}$, 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 choose a national job for any s_l, i.e., $a \geq \min\{e^+(1), \overline{e}\}$, 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 $s_l = \underline{s_l}(f)$ (at $s_l = \overline{s_l}(f)$) from Lemma A4, and increase (decrease) with s_l for $s_l \in (\underline{s_l}(f), s_l^\flat]$ (for $s_l \in [s_l^\sharp, \overline{s_l}(f))$), if $e^+(0) > \frac{\alpha}{1-\alpha}\frac{h_l}{\delta_l}$, $E(e|e<e^+(0)) > \alpha\max\left\{\frac{1}{1-\alpha}\frac{h_l}{\delta_l}, \frac{e^+(0)}{1+\alpha}\right\}$, and $s_l^\flat > \underline{s_l}(f)$, where $s_l^\flat \in (0, s_l^\sharp)$ satisfies $s_l^\flat = (1-\alpha) - \alpha\frac{h_l}{\delta_l E(e|e<e^+(s_l^\flat))}$, 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, δ_n, and δ_l are large, because $e^+(0)$ does not depend on these parameters, $s_l^\flat$ increases with δ_l (since $e^+(s_l)$ increases with δ_l from the proof of Proposition 1 (i)(b)), and $\underline{s_l}(f)$ decreases with A, δ_n, and δ_l and approaches 0 as these parameters increase from Proposition 1 and its proof. The net earnings are maximized at $s_l = 0$ or s_l satisfying $\frac{d(w_n h_n)}{ds_l} = 0$, where the latter 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. From (64) and (65) 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 $\frac{d(w_n h_n)}{ds_l} = 0$) iff $\left(\frac{h_l}{e^+(0)}\right)^{1-\alpha} < \left(\frac{h_l+\delta_l s_l^\flat e^+(s_l^\flat)}{e^+(s_l^\flat)}\right)^{1-\alpha}(1-s_l^\flat)^\alpha$, which holds if

$$\left(\frac{h_l}{e^+(0)}\right)^{1-\alpha}<\left(\frac{h_l+\delta_l s_l^\flat \overline{e}}{\overline{e}}\right)^{1-\alpha}(1-s_l^\flat)^\alpha. \tag{68}$$

The condition holds when δ_l is sufficiently large, because $e^+(0)$ does not depend on A, δ_n, and δ_l, $s_l^\flat$ increases with δ_l, and the RHS increases with $s_l^\flat$:

$$\frac{1-\alpha}{\underline{h}_l+\delta_l s_l^\flat \overline{e}}\delta_l\overline{e} - \frac{\alpha}{1-s_l^\flat} = \frac{\delta_l\overline{e}(1-\alpha-s_l^\flat)-\alpha\underline{h}_l}{(\underline{h}_l+\delta_l s_l^\flat\overline{e})(1-s_l^\flat)} = \alpha\underline{h}_l\frac{\frac{\overline{e}}{E(e|e<e^+(s_l^\flat))}-1}{(\underline{h}_l+\delta_l s_l^\flat\overline{e})(1-s_l^\flat)} > 0. \tag{69}$$

As for workers who 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 $s_l = \underline{s_l}(f)$ (at $s_l = \overline{s_l}(f)$) from Lemma A4, and increase (decrease) with s_l for $s_l \in (\underline{s_l}(f), \min(s_{l,h}^{\triangle}(e), \overline{s_l}(f))]$ (for $s_l \in [s_{l,h}^{\triangledown}(e), \overline{s_l}(f))$ when $s_{l,h}^{\triangledown}(e) < \overline{s_l}(f)$), if $s_{l,h}^{\triangle}(e) > \underline{s_l}(f)$ and $e = a > \max\left\{\alpha\left(\frac{\underline{h}_l}{\delta_l}+e^+(0)\right), \Lambda(\overline{e})\right\}$ from Proposition A1 (ii). The condition holds when A and δ_n are large, because $e^+(0)$, $\Lambda(\overline{e})$, and $s_{l,h}^{\triangle}(e)$ do not depend on these parameters, and $\underline{s_l}(f)$ decreases with A and δ_n. The condition holds when δ_l is large, because $\Lambda(\overline{e})$ decreases with δ_l, $s_{l,h}^{\triangle}(e) \in (0,1)$ from the proof of Proposition A1 (ii) in Appendix D, while $\underline{s_l}(f)$ approaches 0 as δ_l increases, which is from $\underline{s_l}(f)$ being decreasing in δ_l, and from (14) and Lemma A1 in the proof of Proposition 1. Their net earnings are maximized at either $s_l = 0$, or $s_l \in (s_{l,h}^{\triangle}(e), \overline{s_l}(f))$ 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 the maximum.) From (2) and (6), net earnings of such workers at $s_l = 0$ equal

$$(1-\alpha)A\left(\frac{H_n(e^+,0)}{H_l(e^+,0)}\right)^\alpha \underline{h}_l = (1-\alpha)A\left(\delta_n\frac{\alpha}{1-\alpha}\frac{e^+(0)}{\underline{h}_l}\right)^\alpha \underline{h}_l. \tag{70}$$

From (2) and (6), net earnings of such workers with $e = a$ at s_l satisfying $\frac{d(w_l h_l)}{ds_l} = 0$ equal $(1-\alpha)A\left[\frac{\alpha\delta_n(1-s_l)e^+(s_l)}{(1-\alpha)(\underline{h}_l+\delta_l s_l e^+(s_l))}\right]^\alpha(\underline{h}_l+\delta_l s_l e) - e$. The net earnings at such s_l is greater than at $s_l = 0$ iff

$$(1-\alpha)A\left(\delta_n\frac{\alpha}{1-\alpha}\right)^\alpha\left\{\left[\frac{(1-s_l)e^+(s_l)}{\underline{h}_l+\delta_l s_l e^+(s_l)}\right]^\alpha(\underline{h}_l+\delta_l s_l e) - \left(\frac{e^+(0)}{\underline{h}_l}\right)^\alpha \underline{h}_l\right\} - e > 0, \tag{71}$$

which is true if

$$(1-\alpha)A\left(\delta_n\frac{\alpha}{1-\alpha}\right)^\alpha\left\{[(1-s_l)e]^\alpha(\underline{h}_l+\delta_l s_l e)^{1-\alpha} - \left(e^+(0)\right)^\alpha(\underline{h}_l)^{1-\alpha}\right\} - e > 0. \tag{72}$$

When δ_l is sufficiently large, $1-\alpha \in (\underline{s_l}(f), \overline{s_l}(f))$ from Lemma A1 in the proof of Proposition 1. Hence, the above condition holds if

$$(1-\alpha)A\left(\delta_n\frac{\alpha}{1-\alpha}\right)^\alpha\left\{(\alpha e)^\alpha[\underline{h}_l+\delta_l(1-\alpha)e]^{1-\alpha} - \left(e^+(0)\right)^\alpha(\underline{h}_l)^{1-\alpha}\right\} - e > 0, \tag{73}$$

which is true when δ_l is sufficiently large, because the expression inside the curly bracket is large positive ($e^+(0)$ does not depend on δ_l). 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 $s_l = 0$ or $s_l \in (\underline{s_l}(f), \overline{s_l}(f))$ satisfying either $\frac{d(w_n h_n)}{ds_l} = 0$, $\frac{d(w_l h_l)}{ds_l} = 0$, or $a = e = e^+(s_l)$ (s_l at which the switch to a local job occurs). (The earnings at $s_l = \overline{s_l}(f)$ are smaller than the ones when $s_l \to \overline{s_l}(f)$ from above and thus cannot be the maximum.) From the argument above, the net earnings are maximized at the latter s_l when δ_l is sufficiently large. □