Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Oral Skills Pay, Literacy Lags: Differential Wage Returns to Host-Country Language Competencies Among Bangladeshi Migrant Workers in Malaysia

Suchana Das

suchana.das5100@gmail.com

Universiti Kebangsaan Malaysia (UKM)

Hazrul Izuan Bin Shahiri

Universiti Kebangsaan Malaysia (UKM)

Chung Khain Wye

Universiti Kebangsaan Malaysia (UKM)

Research Article

Keywords: Host language competency, RIF decomposition, wage distribution, low-skilled migrant workers, Bangladeshi migrants, occupational task demands

Posted Date: June 22nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9971305/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Title of the Manuscript:

Oral Skills Pay, Literacy Lags: Differential Wage Returns to Host-Country Language Competencies Among Bangladeshi Migrant Workers in Malaysia

Authors
Suchana Das[a]
[a]Faculty of Economics and Management, Universiti Kebangsaan Malaysia (UKM), 43600 UKM, Bangi Selangor, Malaysia
Email: suchana.das5100@gmail.com
ORCiD: https://orcid.org/0009-0005-2783-9839

Hazrul Izuan Bin Shahiri[b]
[b]Faculty of Economics and Management, Universiti Kebangsaan Malaysia (UKM), 43600 UKM, Bangi Selangor, Malaysia
Email: hizuan@ukm.edu.my
ORCiD: https://orcid.org/0000-0002-1344-6074

Wye Chung Khain[c]
[c]Faculty of Economics and Management, Universiti Kebangsaan Malaysia (UKM), 43600 UKM, Bangi Selangor, Malaysia
Email: wyeck@ukm.edu.my
ORCiD: https://orcid.org/0000-0001-9450-942X

Corresponding Author:
Suchana Das[a]
[a]Faculty of Economics and Management, Universiti Kebangsaan Malaysia (UKM), 43600 UKM, Bangi Selangor, Malaysia
Email: suchana.das5100@gmail.com
ORCiD: https://orcid.org/0009-0005-2783-9839

Author Contributions (CRediT Statement)

- Suchana Das: Conceptualization, Methodology, Software, Data curation, Formal analysis, Writing – original draft, Writing – review and editing
- Hazrul Izuan Bin Shahiri: Conceptualization, Supervision, Validation, Writing – review and editing
- Wye Chung Khain: Conceptualization, Supervision, Validation, Writing – review and editing

Oral Skills Pay, Literacy Lags: Differential Wage Returns to Host-Country Language Competencies Among Bangladeshi Migrant Workers in Malaysia

ABSTRACT

Language proficiency is rarely disaggregated in migration economics, yet listening, speaking, reading, and writing competencies may generate unequal wage returns depending on occupational context. This study examines whether these four Malay language skills yield differential wage premiums among low-skilled Bangladeshi migrant workers in Malaysia, and whether returns vary across the wage distribution. Using Recentred Influence Function decomposition and primary survey data from 426 respondents collected via multistage stratified sampling, wage gaps between competent and non-competent users of each skill are estimated and decomposed into explained and unexplained components across the full distribution. The results reveal a consistent oral-literacy divide, where speaking and listening generate statistically significant wage premiums across the distribution, while reading and writing yield smaller and less consistent returns, reflecting the oral task demands of labour-intensive employment. Gaps are most pronounced at the lower and middle quantiles, and the unexplained component dominates at the lower end, indicating that structural factors constrain returns most severely among the most economically vulnerable workers. These findings affirm Human Capital Theory's prediction of positive returns to skill acquisition while supporting Segmented Labour Market Theory in demonstrating that structural conditions attenuate those returns. The study recommends prioritising oral communication training in workforce development programmes targeting low-skilled migrants, alongside structural policy interventions to address wage inequality that skill development alone cannot resolve.

KEYWORDS: *Host language competency; RIF decomposition; wage distribution; low-skilled migrant workers; Bangladeshi migrants; occupational task demands*

JEL codes: *J31, J24, J61, J42, C21*

1. INTRODUCTION

Language proficiency is routinely treated as a single attribute in the migration economics literature, yet the ability to listen, speak, read, and write in a host-country language represents four functionally distinct competencies. In occupational environments organised around verbal instruction, physical task execution, and supervisory communication, the economic returns to oral skills are likely to differ substantially from those to literacy-based skills. Whether oral and literacy skills generate different wage returns, and whether that difference holds across the wage distribution, has rarely been tested. This study addresses that gap among low-skilled Bangladeshi migrant workers in Malaysia, a setting where host language (Malay) functions as the primary language of workplace communication and where the structure of low-skilled employment makes the oral-literacy significant for wage returns.

The broader relationship between host-country language proficiency and labour market outcomes is well established in high-income, migrant-receiving destination country contexts (Chiswick and Miller 1995; Dustmann and Van Soest 2002; Bleakley and Chin 2004). However, this literature predominantly treats language proficiency as a composite measure, even studies that collect data on multiple competencies often collapse them into a single indicator (Liwiński 2019) and is concentrated among skilled or overall migrant populations, leaving middle-income Southeast Asian destinations and the question of skill-specific returns for low-skilled workers substantially underexamined. The theoretical basis for expecting differential returns across skill types is grounded in occupational task demands, where low-skilled employment in labour-intensive sectors is predominantly characterised by manual and communicative tasks rather than literacy-dependent ones (Autor and Handel 2013), generating a prediction that oral competencies should yield larger wage premiums than literacy skills. The limited direct evidence on skill-specific returns, principally from Dustmann (1994), points in

this direction but remains confined to high-income economies, leaving the question unresolved for the present context of labour migration.

This study draws on Human Capital Theory (Becker 1964; Mincer 1974), which predicts positive returns to language skill acquisition, and Segmented Labour Market Theory (Piore 1979; Doeringer and Piore 1971), which anticipates that structural conditions will reduce those returns among workers in the secondary labour market segment. Together these perspectives raise the question of whether language premiums are structurally constrained in ways that vary across the wage distribution. This study examines the wage returns to four host language (Malay) competencies, namely listening, speaking, reading, and writing, among low-skilled Bangladeshi migrant workers in Selangor, Malaysia, using the Recentred Influence Function decomposition approach.

Three research questions of this study are, first, do oral competencies generate larger wage premiums than literacy competencies? Second, how do wage gaps vary across the wage distribution for each skill category? Third, to what extent do explained and unexplained factors account for wage disparities at the lower quantiles?

This study makes three contributions. First, rather than treating language proficiency as a single composite measure, it disaggregates it into four distinct competencies, namely, listening, speaking, reading, and writing, allowing their wage returns to be examined separately. Second, it applies RIF decomposition across the full wage distribution, making it possible to identify where structural barriers to language returns are most binding rather than relying on mean-based estimates that can mask this variation. Third, it extends the skill-specific language returns literature, previously confined largely to high-income destinations, to a middle-income economy where the structure of low-skilled employment differs substantially from the contexts in which this literature has developed.

2. LANGUAGE SKILLS AND WAGE RETURNS AMONG MIGRANT WORKERS

Human Capital Theory (Becker 1964; Mincer 1974) suggests that individuals invest in skill acquisition to enhance their productive capacity and earnings potential. Within this framework, host-country language proficiency constitutes a form of human capital investment that improves workplace communication, broadens job access, and signals adaptability to employers, all of which translate into higher wages. The theory predicts that greater language competency should translate directly into higher labour market returns (Chiswick and Miller 1995; Bleakley and Chin 2004).

However, Human Capital Theory assumes largely competitive labour markets where returns to skill reflect individual productivity, an assumption that does not fit well with the position of low-skilled migrant workers. Segmented Labour Market Theory (Piore 1979; Doeringer and Piore 1971) offers an appropriate justification, distinguishing between a primary segment characterised by stable employment and higher wages, and a secondary segment characterised by job instability, low wages, and limited mobility. Low-skilled migrant workers are disproportionately concentrated in this secondary segment, where wages are governed less by individual productivity than by institutional practices and the dynamics of labour supply; persistently high inflows of low-skilled foreign labour further suppress wages and weaken individual bargaining power (Borjas 2003). The two theories agree that language skills can raise wages, but they differ on the factors that determine the size of that gain. Human Capital Theory traces it to the productivity gains the skill itself produces, while Segmented Labour Market Theory points to the structural and institutional conditions that limit the extent to which that gain shows up in wages, particularly at the lower end of the wage distribution.

The positive relationship between host-country language proficiency and migrant earnings is well established (Chiswick and Miller 1995; Dustmann and Van Soest 2002), with causal

identification provided through variation in age at arrival (Bleakley and Chin 2004). Fewer studies, however, have examined whether different language skills carry different returns, and most that collect data on multiple competencies still reduce them to a single measure. Liwiński (2019), for instance, gathered self-assessed ratings for speaking, writing, reading, and listening but used only the speaking measure in his wage analysis, citing concerns about measurement bias in the others. The exception is Dustmann (1994), who found that speaking fluency contributed significantly to migrant earnings in Germany, while returns to writing varied by occupational category, with oral skills carrying more weight in manual and trade occupations. This finding suggests that the wage value of a language skill is shaped less by the skill itself than by what the job actually requires the worker to do. This is a pattern that should be especially pronounced among low-skilled migrant workers, whose jobs are concentrated in manual and service occupations where oral communication, rather than literacy, dominates daily task demands.

A separate but related body of work has examined how returns to individual characteristics vary across the wage distribution, recognising that mean estimates may mask substantial heterogeneity. The RIF decomposition method extends the Oaxaca-Blinder framework to unconditional quantiles, allowing estimation of the total wage gap and its explained and unexplained components at any point of the distribution (Firpo, Fortin and Lemieux 2009). Studies have found that the unexplained component is largest at lower quantiles, consistent with structural and unobserved influences bearing most heavily on lower-income workers (Rios-Avila 2020). Distributional methods have also been applied to migrant wage gaps in high-income contexts, where immigration effects on native wages have been shown to concentrate at the lower end of the distribution (Dustmann, Frattini and Preston 2013), but they have not yet been applied to skill-specific language returns.

The existing literature converges on a common limitation that language returns have been examined largely at the mean, in high-income host country settings, using measures that do not separate listening, speaking, reading, and writing as distinct competencies. This study addresses that limitation directly, estimating separate wage returns to four Malay language competencies among low-skilled Bangladeshi migrant workers in Malaysia and applying RIF decomposition to identify how those returns differ across the wage distribution.

3. METHODOLOGY

3.1. Data and Sample

The data used in this study were collected through a structured survey of low-skilled Bangladeshi migrant workers employed in Malaysia. The survey instrument was designed by the author, reviewed and validated by academic experts, and administered through face-to-face data collection in Selangor, Malaysia. The study received ethical approval from UKM Research Ethics Secretariat on 11 July 2025, under the application reference JEP-2024-880 (Letter Reference: UKM PPI/111/8) prior to data collection. A stratified sampling procedure was used to ensure that the sample represented variation across both industrial sectors of employment and regional origin within Bangladesh. The target population was divided by industry of employment (construction, manufacturing, and service sectors) and regional origin within Bangladesh (metropolitan versus non-metropolitan), worksites across Selangor were selected proportionally across these strata, and workers were recruited on-site and at community gathering points. The final sample consists of 426 workers employed in low-skilled construction, manufacturing, or service occupations at the time of data collection.

3.2. Measurement of Variables

The dependent variable in all estimations is the natural logarithm of net monthly wages in Malaysian Ringgit, following standard practice in the wage determination literature (Mincer 1974). Malay language competency was assessed across four domains, namely, listening, speaking, reading, and writing, using a six-point self-reported scale adapted from the Common European Framework of Reference for Languages descriptors (Council of Europe 2001). Respondents scoring four or above, corresponding to the independent user level, were classified as competent; those scoring below four were classified as non-competent. Each domain is retained as a separate binary variable, allowing wage returns to be estimated and compared across the four competencies individually rather than through a single measurement.

The estimations include a set of control variables. Education is measured as years of formal schooling, and relevant work experience is entered in years together with its quadratic term (Mincer 1974). Marital status is a dichotomous variable equal to one if the worker is married, and number of dependent family members is included as a continuous measure of financial obligations. Lastly, migration cost, the natural logarithm of initial migration expenditure in Malaysian Ringgit.

3.3. Empirical Framework

Based on the data collected this study develops the following baseline empirical estimation for analysis,

$$lnwage_i = \alpha_0 + \beta_1 X_i + \mu_i \qquad \text{Eq. (1)}$$

$lnwage_i$ denotes the natural logarithm of the monthly wage of worker *i* (measured in RM), X_i represents a vector of control variables comprising education, experience (along with its

squared term), marital status, family size, and migration cost, α_0 is the constant term, and μ_i is the stochastic error term.

This study employs the Recentred Influence Function (RIF) decomposition of Firpo, Fortin, and Lemieux (2009). RIF decomposition decomposes unconditional wage distribution, allowing examination of how the wage gap between competent and non-competent users varies at multiple quantiles of the wage distribution. This distributional flexibility is particularly important given the theoretical expectation, grounded in Segmented Labour Market Theory, that structural constraints bear most heavily on workers at the lower end of the distribution.

The RIF of an unconditional quantile is defined as:

$$RIF(y; q_\tau) = q_\tau + \frac{\tau - 1(y \leq q_\tau)}{f_Y(q_\tau)} \qquad \text{Eq. (2)}$$

where q_τ is the τ-th quantile of the unconditional wage distribution F_Y, $f_Y(q_\tau)$ is the density of the wage distribution evaluated at q_τ, and $1(y \leq q_\tau)$ is an indicator function equal to one if the observed wage y is at or below the quantile. The RIF values are subsequently regressed on the vector of covariates using ordinary least squares:

$$E[RIF(y; q_\tau) \mid \mathrm{X}] = \mathrm{X}\beta^\tau \qquad \text{Eq. (3)}$$

where β^τ denotes the vector of RIF regression coefficients at quantile τ. The total wage gap at quantile τ between competent users (group C) and non-competent users (group N) is then decomposed following the Oaxaca-Blinder framework applied to the RIF regression coefficients:

$$\hat{\Delta}^\tau = \underbrace{(\bar{\mathrm{X}}_C - \bar{\mathrm{X}}_N)\hat{\beta}_N^\tau}_{\text{Explained}} + \underbrace{\bar{\mathrm{X}}_C\left(\hat{\beta}_C^\tau - \hat{\beta}_N^\tau\right)}_{\text{Unexplained}} \qquad \text{Eq. (4)}$$

where $\bar{\mathrm{X}}_C$ and $\bar{\mathrm{X}}_N$ are the vectors of mean characteristics for the competent and non-competent groups respectively, and $\hat{\beta}_C^\tau$ and $\hat{\beta}_N^\tau$ are the corresponding RIF regression coefficients estimated separately for each group at quantile τ. The explained component captures the portion of the wage gap attributable to differences in observable characteristics such as education and work

experience; the unexplained component captures differences in returns to those characteristics and unobservable, which this paper interprets, in part, as reflecting the structural constraints anticipated by Segmented Labour Market Theory.

The decomposition is estimated separately for each of the four language competency variables at the mean and at the 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} quantiles of the unconditional wage distribution, using the "*rifoaxaca"* command in Stata as documented by Rios-Avila (2020). For writing competency, estimation at the 10^{th} quantile was not achievable due to convergence constraints; decomposition results for writing are therefore reported at the mean and at the 25^{th}, 50^{th}, 75^{th}, and 90^{th} quantiles only.

3.4. Descriptive Statistics

Table 1 presents descriptive statistics for the full sample and by competency group across all four skill domains. Competent speakers earn a mean log wage of 7.75 compared to 7.57 among non-competent speakers, and competent listeners earn 7.70 compared to 7.52 among non-competent listeners. The gaps are smaller for the literacy skills, with competent readers earning 7.78 against 7.64 and competent writers earning 7.79 against 7.64. Across all four competency splits, competent workers have more years of education and work experience than their non-competent counterparts. The descriptive statistics thus provide preliminary evidence of a positive wage gap associated with Malay language competency across all four skills, with the gap appearing larger for the oral skills (speaking and listening) than for reading and writing.

4. RESULTS AND DISCUSSION

The RIF decomposition results across all four Malay language competencies consistently show that competent users earn higher wages than non-competent users across most points of the wage distribution. Positive and statistically significant wage gaps are observed across all

quantiles for speaking and listening, while reading and writing yield positive gaps that are less consistently significant. Mean wage gaps range from 0.147 log points for reading to 0.178 log points for speaking, with listening (0.174) closely following. Across all four skills, wage gaps are largest at the lower and middle quantiles and narrow progressively toward the upper end of the distribution, a pattern that carries direct implications for understanding how structural conditions shape the returns to language skills among this group of workers.

Grouping the four skills by type reveals a consistent oral-literacy split. The two oral competencies, speaking and listening, generate larger mean wage premiums (0.178 and 0.174 log points respectively) and statistically significant gaps across the full distribution. The two literacy competencies, reading and writing, yield smaller mean premiums (0.147 and 0.154 log points respectively) with gaps that are less consistently significant, particularly at the median. It reflects the structure of communicative task demands in low-skilled occupational environments, where verbal instruction, supervisory coordination, and spoken task management constitute the primary modes of workplace interaction, while written communication plays a limited role in daily work performance (Autor and Handel 2013). The results provide direct empirical support for the prediction that oral competencies carry larger wage returns than literacy skills in this occupational context.

4.1. Speaking Competency and Wage Returns

Table 2 presents the RIF decomposition results for Malay speaking competency. Competent speakers earn a mean log wage of 7.746, equivalent to RM 2,312.30, compared to 7.568, equivalent to RM 1,935.27, for non-competent speakers, yielding the largest mean wage gap among the four competencies at 0.178 log points. The gap is statistically significant throughout the distribution, reaching its highest values at the 25th quantile (0.214) and the 10th quantile (0.202), and narrowing to 0.129 at the 75th quantile. The consistency and magnitude of the

speaking wage premium across the distribution provide strong empirical support for the argument that oral communication ability in Malay is the most economically consequential language skill for this group of workers.

The decomposition results reveal a pronounced distributional pattern in the sources of the wage gap. At the 10th and 25th quantiles, the unexplained component accounts for 73.76 percent and 63.08 percent of the total gap respectively, pointing to the influence of structural and unobserved factors among lower earners. At the 75th quantile, the explained component rises sharply to 95.35 percent, indicating that among higher earners the speaking wage premium is almost entirely attributable to differences in observable human capital characteristics. This contrast between the lower and upper ends of the distribution is the most pronounced among all four competencies.

4.2. Listening Competency and Wage Returns

Table 3 presents the RIF decomposition results for Malay listening competency. Competent listeners earn a mean log wage of 7.698, equivalent to RM 2,203.94, compared to 7.525, equivalent to RM 1,853.81, for non-competent listeners, yielding a mean wage gap of 0.174 log points. The gap is statistically significant across all quantiles, ranging from 0.145 at the 10th and 90th quantiles to 0.209 at the median. The concentration of the largest gaps at the 25th and 50th quantiles suggests that the wage premium for listening competency is most pronounced among workers in the lower-to-middle range of the wage distribution.

The decomposition reveals a distributional pattern consistent with that observed for speaking. At the lower quantiles, the unexplained component dominates, accounting for 77.93 percent of the total gap at the 10th quantile and 73.08 percent at the 25th quantile, indicating that among the lowest earners, wage differences between competent and non-competent listeners cannot be explained primarily by observable human capital characteristics. At the 75th quantile,

however, the explained component accounts for 87.88 percent of the gap, suggesting that among higher earners, differences in education, experience, and other measurable characteristics largely account for the wage differential. This shift in the relative contributions of the explained and unexplained components across the distribution is theoretically consistent with Segmented Labour Market Theory, which predicts that structural constraints position heavily on workers at the lower end of the wage distribution (Piore 1979).

4.3. Reading Competency and Wage Returns

The very low prevalence of reading competency in the sample, with only 18 of 426 respondents classified as competent (4.23%), is itself a practical finding. It reflects the structural reality that literacy skills in Malay are rarely acquired by low-skilled Bangladeshi migrant workers in this occupational context, which constrains the wage variation the decomposition can detect and accounts for the wider confidence intervals observed at some quantiles.

Table 4 presents the RIF decomposition results for Malay reading competency. Competent readers earn a mean log wage of 7.783, equivalent to RM 2,399.46, compared to 7.636, equivalent to RM 2,071.44, for non-competent readers, producing a mean wage gap of 0.147 log points, the smallest mean gap among the four competencies. The gap is largest at the 25th quantile (0.169) and 10th quantile (0.165) and narrows substantially at the 75th quantile (0.067). The wage gap at the median is positive but does not reach conventional levels of statistical significance, indicating that the earnings differential between competent and non-competent readers is less consistently established at the middle of the distribution.

The decomposition follows a similar distributional pattern to the oral competencies, with the unexplained component dominating at the lower quantiles. At the 10th quantile, 72.73 percent of the total gap is unexplained, declining to 26.87 percent at the 75th quantile. The relatively modest and less consistently significant mean wage gap for reading, combined with the very

high proportion of respondents classified as low proficiency in this skill at 95.77 percent, indicates that reading competency in Malay is both rare and inconsistently rewarded in this labour market segment. The absence of reading proficiency in this sample limits the wage variation the decomposition can detect. The estimated effects therefore reflect a lack of competency variation in the sample rather than evidence that reading carries no wage value in this labour market.

4.4. Writing Competency and Wage Returns

Table 5 presents the RIF decomposition results for Malay writing competency. The extremely low prevalence of writing competency, with only 9 respondents classified as competent (2.11%), reflects the near-complete absence of written Malay exposure in low-skilled occupational environments. This distributional reality, rather than a limitation of the estimation strategy, explains the reduced precision of the decomposition results and reinforces the paper's central argument that literacy skills remain economically and practically marginal in this labour market segment. Due to estimation constraints arising from the very low proportion of competent writers in the sample, results are reported at the mean and at the 25th, 50th, 75th, and 90th quantiles only. Competent writers earn a mean log wage of 7.793, equivalent to RM 2,423.58, compared to 7.639, equivalent to RM 2,077.67, for non-competent writers, producing a mean wage gap of 0.154 log points. The largest positive differences in wages appear at the median (0.195) and the 25th quantile (0.168), while the gap narrows considerably at the 75th quantile (0.059), the smallest observed gap across all skills and quantiles in this study.

The unexplained component is the dominant contributor to the writing wage gap across the distribution, accounting for 74.03 percent at the mean, 76.19 percent at the 25th quantile, and 75.90 percent at the median. While the wage gaps for writing are positive across all estimated quantiles, they do not reach conventional levels of statistical significance beyond the mean,

where significance is observed at the 10 percent level only. This pattern of weak statistical significance is consistent with the descriptive evidence showing that 97.89 percent of respondents are classified as low proficiency in writing, suggesting that written Malay competency is neither widely acquired nor consistently rewarded in this labour market segment. This reinforces the broader argument that literacy-based competencies remain economically marginal for low-skilled migrant workers whose occupational exposure to Malay language is predominantly oral.

4.5. Discussion

Three consistent patterns emerge from the results across the four competency skills, each carrying distinct theoretical implications.

The first and most directly relevant pattern concerns the oral-literacy divide in wage returns. Speaking and listening generate the largest premiums at the mean and yield statistically significant gaps throughout the distribution, while reading and writing produce smaller and less consistently significant returns. This ordering aligns with the task-demand structure of low-skilled employment identified by Autor and Handel (2013), in which construction and manufacturing occupations are predominantly characterised by manual and communicative tasks rather than literacy-dependent ones. Workers in these sectors are rewarded for the language skills most directly applicable to their productive roles, and in occupational environments organised around verbal instruction and supervisory communication, those skills are oral. The absence of reading and writing proficiency in this sample, at 95.77 percent and 97.89 percent below the functional threshold respectively, further constrains the wage variation attributable to literacy competencies. These findings extend the skill-specific evidence of Dustmann (1994) and Liwiński (2019) to a middle-income Southeast Asian context and

demonstrate that the oral-literacy divide in language returns is present and arguably more pronounced where the occupational range available to workers is narrower.

The second pattern concerns the concentration of wage gaps at the lower and middle quantiles of the distribution. Across all four competencies, the largest absolute gaps appear at the 25th and 50th quantiles and narrow progressively toward the upper end. Workers at the lower and middle quantiles are precisely those whose employment conditions are most sensitive to labour market entry barriers, informal wage-setting arrangements, and limited access to employer networks. Research on language and labour market performance suggests that host-country language skills mediate access to employment opportunities, with language barriers restricting workers to a narrower range of job options and limiting their ability to respond to wage differentials across employers (Dustmann and Fabbri 2003). In this context, the wage gap between competent and non-competent users is widest where language disadvantage most constrains labour market participation, and narrows at higher quantiles where other factors, including seniority and sector-specific experience, may partially offset language disadvantage.

The third pattern concerns the dominant contribution of the unexplained component at the lower quantiles. Across all four skills, structural and unobserved factors account for the majority of the wage gap among the lowest earners, while differences in observable human capital characteristics explain a considerably larger share at the upper end of the distribution. This distributional shift is theoretically consistent with Segmented Labour Market Theory. Workers at the lower end of the distribution are concentrated in the secondary labour market segment, where wages are determined less by individual productive characteristics than by institutional practices, employer's wage setting power, and the structural conditions of labour supply (Piore 1979; Doeringer and Piore 1971). The sustained high inflow of low-skilled foreign labour into the Malaysian economy suppresses wage floors and systematically weakens the bargaining position of workers in this segment (Borjas 2003; Devadason 2021), limiting

the wage returns that even language-competent workers can realise at the lower end of the distribution. These findings are consistent with Human Capital Theory in confirming positive returns to language skill acquisition, while the distributional patterns provide empirical support for Segmented Labour Market Theory in demonstrating that structural conditions attenuate those returns most sharply at the lower end of the distribution.

These findings carry direct implications for workforce development policy. The consistently larger and more statistically robust wage premiums for speaking and listening indicate that oral competency training should be prioritised in workforce development programmes targeting low-skilled migrant workers, designed specifically around the supervisory and peer communication demands of construction and manufacturing environments. Equally important is what the findings imply about where training resources should not be directed. Given that 95.77 percent of respondents fall below the functional threshold in reading and 97.89 percent in writing, literacy-focused language programmes would reach very few workers currently positioned to convert that skill into wage gains, and oral communication training offers a more tractable near-term priority. Pre-departure and in-country language programmes in both Bangladesh and Malaysia should therefore prioritise oral communication content calibrated to workplace contexts. At the structural level, the persistent dominance of the unexplained wage gap at the lower quantiles indicates that skill development alone is insufficient to close wage disparities among the most vulnerable workers. Addressing this dimension of wage inequality requires stronger enforcement of minimum wage legislation, greater transparency in wage-setting practices, and institutional mechanisms that strengthen the bargaining position of low-skilled foreign workers in the Malaysian labour market.

5. CONCLUSION

This study examined whether distinct Malay language competencies generate differential wage returns among low-skilled Bangladeshi migrant workers in Selangor, Malaysia, and whether those returns vary across the wage distribution. The RIF decomposition results confirm both. Oral competencies generate larger and more consistent wage premiums than literacy-based skills, reflecting the communicative task demands of low-skilled employment. Wage gaps are most pronounced at the lower and middle quantiles, and the unexplained component dominates at the lower end of the distribution across all four skills. Together, these findings confirm the predictions of Human Capital Theory regarding positive returns to language skill acquisition while demonstrating, consistent with Segmented Labour Market Theory, that structural conditions attenuate those returns most sharply at the lower end of the distribution.

Two limitations bear on the interpretation of these findings. Malay language competency is self-reported and classified at a fixed binary threshold, which may introduce measurement error and sensitivity to the chosen cutoff. The cross-sectional design and restriction to Selangor constrain causal inference and generalisability. Longitudinal data, objective language assessments, and broader geographic coverage would strengthen future work in this area.

DECLARATIONS

a. Acknowledgements

The authors have no acknowledgements to declare.

b. Data Availability Statement

This study is based on primary survey data collected from low-skilled Bangladeshi migrant workers in Malaysia. The data are not publicly available as they were collected under conditions of participant confidentiality. The dataset and Stata estimation files used in this study are available from the corresponding author upon reasonable request.

c. Conflict of Interest

The authors declare that there are no conflicts of interest related to the publication of this manuscript.

d. Permission

No copyrighted material requiring permission has been used in this study.

e. Ethics Approval

This study involved the collection of primary survey data from human participants. All respondents participated voluntarily and were informed of the purpose of the study prior to data collection. No sensitive personal information was collected and all responses were anonymised. Formal ethical approval was obtained in accordance with the requirements of the authors' institutional affiliation.

f. Funding

No financial support was received for the conduct of this research or the preparation of this article.

LIST OF TABLES

Table 1: Descriptive Statistics by Malay Language Competency

Variable	Speaking		Listening		Reading		Writing	
	NC (n=248)	C (n=178)	NC (n=137)	C (n=289)	NC (n=408)	C (n=18)	NC (n=417)	C (n=9)
Log monthly wage	7.57 (0.23)	7.75 (0.19)	7.52 (0.23)	7.70 (0.21)	7.64 (0.23)	7.78 (0.21)	7.64 (0.23)	7.79 (0.19)
Education (years)	7.72 (3.18)	9.01 (3.89)	7.45 (3.33)	8.64 (3.59)	8.08 (3.44)	12.33 (3.56)	8.17 (3.51)	12.44 (2.79)
Work experience (years)	5.70 (5.53)	8.33 (4.23)	4.92 (5.85)	7.69 (4.59)	6.80 (5.27)	6.89 (2.63)	6.82 (5.22)	5.78 (2.95)
Marital status	0.43	0.63	0.43	0.55	0.51	0.67	0.51	0.78
Number of dependants	3.31 (1.20)	3.84 (1.16)	3.34 (1.36)	3.62 (1.13)	3.52 (1.22)	3.72 (0.89)	3.52 (1.22)	4.33 (0.71)
Log migration cost	9.57 (0.37)	9.40 (0.31)	9.64 (0.37)	9.43 (0.32)	9.50 (0.36)	9.39 (0.29)	9.50 (0.36)	9.48 (0.30)

Notes: NC = Non-Competent; C = Competent. Competency groups are defined separately for each skill domain. Continuous variables are reported as mean (standard deviation); binary variables are reported as proportion. Log monthly wage is the natural logarithm of net monthly wages in Malaysian Ringgit. Education is measured in years of formal schooling. Work experience is measured in years. Marital status is a binary variable equal to one if the worker is married. Number of dependants is the count of dependent family members. Log migration cost is the natural logarithm of initial migration expenditure in Malaysian Ringgit. Competency is defined as a score of four or above on a six-point self-reported proficiency scale adapted from the Common European Framework of Reference for Languages descriptors, corresponding to the independent user level. Source: Author's estimation using primary survey data.

Table 2 - Decomposition of Wage Differentials by Speaking Competency in Malay Language: Evidence from RIF–FFL Estimations

	RIF Mean	RIF Q (10)	RIF Q (25)	RIF Q (50)/ Median	RIF Q (75)	RIF Q (90)
Mean $lnwage_i$ of	7.746***	7.531***	7.630***	7.773***	7.893***	8.031***
Group_1 (C)	(RM 2,312.30)	(RM 1,864.97)	(RM 2,059.05)	(RM 2,375.59)	(RM 2,678.47)	(RM 3,074.82)
Mean $lnwage_i$ of	7.568***	7.329***	7.416***	7.566***	7.765***	7.882***
Group_2 (N)	(RM 1,935.27)	(RM 1,523.86)	(RM 1,662.37)	(RM 1,931.40)	(RM 2,356.66)	(RM 2,649.17)
Wage Gap	0.178***	0.202***	0.214***	0.207***	0.129***	0.149***
Explained	0.094*** (52.81%)	0.053*** (26.24%)	0.079*** (36.92%)	0.129*** (62.32%)	0.123*** (95.35%)	0.044 (29.53%)
Unexplained	0.084*** (47.19%)	0.149*** (73.76%)	0.135*** (63.08%)	0.078** (37.68%)	0.006 (4.65%)	0.105** (70.47%)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Table 2: The dependent variable is log of monthly wages. RIF splits the total wage gap into explained and unexplained components. Percentages in parentheses indicate the share of the total gap; Table 2 is the First Block / Overall Output from RIF Wage Decomposition; Complete Estimations provided in *Appendix A;* Source: Author's estimation using primary data collected

Table 3 - Decomposition of Wage Differentials by Listening Competency in Malay Language: Evidence from RIF–FFL Estimations

	RIF Mean	RIF Q (10)	RIF Q (25)	RIF Q (50)/ Median	RIF Q (75)	RIF Q (90)
Mean $lnwage_i$ of	7.698***	7.462***	7.555***	7.734***	7.889***	8.011***
Group_1 (C)	(RM 2,203.94)	(RM 1,740.63)	(RM 1,910.27)	(RM 2,284.72)	(RM 2,667.77)	(RM 3,013.93)
Mean $lnwage_i$ of	7.525***	7.316***	7.347***	7.525***	7.724***	7.866***
Group_2 (N)	(RM 1,853.81)	(RM 1,504.18)	(RM 1,551.53)	(RM 1,853.81)	(RM 2,261.99)	(RM 2,607.12)
Wage Gap	0.174***	0.145***	0.208***	0.209***	0.165***	0.145***
Explained	0.080*** (45.98%)	0.032** (22.07%)	0.056*** (26.92%)	0.090*** (43.06%)	0.145*** (87.88%)	0.047 (32.41%)
Unexplained	0.094*** (54.02%)	0.113*** (77.93%)	0.152*** (73.08%)	0.119*** (56.94%)	0.020 (12.12%)	0.098 (67.59%)

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Notes Table 3: The dependent variable is log of monthly wages. RIF splits the total wage gap into explained and unexplained components. Percentages in parentheses indicate the share of the total gap; Table 3 is the First Block / Overall Output from RIF Wage Decomposition; Complete Estimations provided in *Appendix B;* Source: Author's estimation using primary data collected

Table 4 - Decomposition of Wage Differentials by Reading Competency in Malay Language: Evidence from RIF–FFL Estimations

	RIF Mean	RIF Q (10)	RIF Q (25)	RIF Q (50)/ Median	RIF Q (75)	RIF Q (90)
Mean $lnwage_i$ of	7.783***	7.542***	7.666***	7.804***	7.917***	8.085***
Group_1 (C)	(RM 2,399.46)	(RM 1,885.60)	(RM 2,134.53)	(RM 2,450.38)	(RM 2,743.53)	(RM 3,245.42)
Mean $lnwage_i$ of	7.636***	7.378***	7.497***	7.675***	7.850***	7.949***
Group_2 (N)	(RM 2,071.44)	(RM 1,600.39)	(RM 1,802.63)	(RM 2,153.82)	(RM 2,565.73)	(RM 2,832.74)
Wage Gap	0.147**	0.165**	0.169**	0.129	0.067	0.136
Explained	0.061** (41.50%)	0.045**(27.27%)	0.079** (46.75%)	0.086** (66.67%)	0.049 (73.13%)	0.052* (38.24%)
Unexplained	0.086 (58.50%)	0.120* (72.73%)	0.090 (53.25%)	0.043 (33.33%)	0.018 (26.87%)	0.084 (61.76%)

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Notes Table 4: The dependent variable is log of monthly wages. RIF splits the total wage gap into explained and unexplained components. Percentages in parentheses indicate the share of the total gap; Table 4 is the First Block / Overall Output from RIF Wage Decomposition; Complete Estimations provided in *Appendix C;* Source: Author's estimation using primary data collected

Table 5 - Decomposition of Wage Differentials by Writing Competency in Malay Language: Evidence from RIF–FFL Estimations

	RIF Mean	RIF Q (25)	RIF Q (50)/ Median	RIF Q (75)	RIF Q (90)
Mean $lnwage_i$ of	7.793***	7.668***	7.873***	7.913***	8.083***
Group_1 (C)	(RM 2,423.58)	(RM 2,138.80)	(RM 2,625.43)	(RM 2,732.58)	(RM 3,238.94)
Mean $lnwage_i$ of	7.639***	7.500***	7.677***	7.854***	7.952***
Group_2 (N)	(RM 2,077.67)	(RM 1,808.04)	(RM 2,158.14)	(RM 2576.02)	(RM 2,841.25)
Wage Gap	0.154*	0.168	0.195	0.059	0.131
Explained	0.040 (25.97%)	0.040 (23.81%)	0.048 (24.62%)	0.038 (64.41%)	0.030 (22.90%)
Unexplained	0.114 (74.03%)	0.128 (76.19%)	0.148 (75.90%)	0.021 (35.59%)	0.102 (77.86%)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; Table 5: The dependent variable is log of monthly wages. RIF splits the total wage gap into explained and unexplained components. Percentages in parentheses indicate the share of the total gap; Table 5 is the First Block / Overall Output from RIF Wage Decomposition; Complete Estimations provided in *Appendix D;* Source: Author's estimation using primary data collected

APPENDIX A
DISTRIBUTIONAL WAGE DECOMPOSITION (SPEAKING)

	DV- Log Wage	RIF Mean	RIF Q (10)	RIF Q (25)	RIF Q (50)/ Median	RIF Q (75)	RIF Q (90)
Overall	Group_1 Competent	7.746*** (0.014)	7.531*** (0.014)	7.630*** (0.021)	7.773*** (0.021)	7.893*** (0.020)	8.031*** (0.026)
	Group_2 Non-Competent	7.568*** (0.015)	7.329*** (0.014)	7.416*** (0.017)	7.566*** (0.022)	7.765*** (0.024)	7.882*** (0.023)
	Difference	0.178*** (0.020)	0.202*** (0.020)	0.214*** (0.027)	0.207*** (0.030)	0.129*** (0.031)	0.149*** (0.035)
	Explained	0.094*** (0.017)	0.053*** (0.016)	0.079*** (0.019)	0.129*** (0.024)	0.123*** (0.029)	0.044 (0.028)
	Unexplained	0.084*** (0.024)	0.149*** (0.018)	0.135*** (0.027)	0.078** (0.033)	0.006 (0.043)	0.105** (0.048)
Explained	Education	0.004 (0.006)	0.002 (0.005)	-0.000 (0.007)	0.006 (0.009)	0.001 (0.010)	-0.010 (0.011)
	Experience	0.098*** (0.023)	0.054*** (0.019)	0.090*** (0.025)	0.148*** (0.035)	0.142*** (0.036)	0.024 (0.027)
	Experience Sq	-0.021* (0.011)	-0.016* (0.009)	-0.025** (0.012)	-0.034** (0.017)	-0.026* (0.015)	0.006 (0.011)
	Marital Status	0.002 (0.006)	0.009* (0.005)	-0.000 (0.007)	0.002 (0.009)	-0.007 (0.010)	0.013 (0.011)
	No. of Family	0.010 (0.006)	0.005 (0.005)	0.015* (0.008)	0.008 (0.009)	0.012 (0.011)	0.010 (0.012)
	Log Migration Cost	0.000 (0.001)	-0.001 (0.002)	-0.001 (0.002)	0.000 (0.002)	0.001 (0.002)	0.001 (0.002)
Unexplained	Education	0.060 (0.053)	0.058 (0.057)	0.097 (0.064)	0.056 (0.077)	-0.006 (0.088)	0.169* (0.099)
	Experience	-0.004 (0.102)	-0.125 (0.115)	0.372** (0.174)	-0.201 (0.151)	-0.305** (0.147)	0.236 (0.198)
	Experience Sq	0.003 (0.053)	0.070 (0.053)	-0.156* (0.084)	0.111 (0.077)	0.109 (0.080)	-0.101 (0.114)
	Marital Status	-0.032 (0.026)	-0.041 (0.027)	-0.013 (0.037)	-0.031 (0.039)	-0.024 (0.041)	-0.097* (0.051)
	No. of Family	0.027 (0.060)	0.002 (0.058)	0.036 (0.080)	0.145* (0.086)	0.002 (0.101)	-0.139 (0.130)
	Log Migration Cost	0.056 (0.509)	0.803 (0.629)	0.844 (0.775)	-0.084 (0.724)	-0.925 (0.716)	-0.994 (0.844)
	_cons	-0.024 (0.521)	-0.618 (0.644)	-1.044 (0.788)	0.081 (0.737)	1.153 (0.726)	1.031 (0.853)
N	*Group 1*				178		
	Group 2				248		

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

APPENDIX B
DISTRIBUTIONAL WAGE DECOMPOSITION (LISTENING)

	DV- Log Wage	RIF Mean	RIF Q (10)	RIF Q (25)	RIF Q (50)/ Median	RIF Q (75)	RIF Q (90)
Overall	Group_1 Competent	7.698*** (0.012)	7.462*** (0.015)	7.555*** (0.017)	7.734*** (0.020)	7.889*** (0.017)	8.011*** (0.024)
	Group_2 Non-Competent	7.525*** (0.020)	7.316*** (0.016)	7.347*** (0.022)	7.525*** (0.027)	7.724*** (0.039)	7.866*** (0.034)
	Difference	0.174*** (0.023)	0.145*** (0.022)	0.208*** (0.028)	0.209*** (0.034)	0.165*** (0.043)	0.145*** (0.041)
	Explained	0.080*** (0.022)	0.032** (0.016)	0.056*** (0.021)	0.090*** (0.028)	0.145*** (0.046)	0.047 (0.043)
	Unexplained	0.094*** (0.028)	0.113*** (0.022)	0.152*** (0.030)	0.119*** (0.037)	0.020 (0.060)	0.098 (0.065)
Explained	Education	0.002 (0.007)	0.006 (0.005)	0.000 (0.007)	-0.004 (0.010)	-0.018 (0.017)	-0.008 (0.015)
	Experience	0.079*** (0.026)	0.023 (0.017)	0.056** (0.025)	0.101*** (0.036)	0.179*** (0.059)	0.039 (0.049)
	Experience Sq	-0.012 (0.010)	-0.004 (0.006)	-0.013 (0.010)	-0.017 (0.014)	-0.034 (0.025)	0.001 (0.015)
	Marital Status	0.008 (0.006)	0.009* (0.005)	0.007 (0.006)	0.014 (0.009)	0.008 (0.011)	0.006 (0.010)
	No. of Family	0.003 (0.004)	-0.001 (0.003)	0.004 (0.005)	-0.004 (0.006)	0.009 (0.010)	0.008 (0.010)
	Log Migration Cost	0.000 (0.000)	-0.000 (0.002)	-0.000 (0.000)	0.000 (0.001)	0.000 (0.003)	0.000 (0.001)
Unexplained	Education	0.073 (0.059)	0.009 (0.048)	0.087 (0.063)	0.161** (0.081)	0.199 (0.122)	0.164 (0.123)
	Experience	0.095 (0.081)	0.187* (0.102)	0.259** (0.109)	0.141 (0.119)	-0.281* (0.151)	0.280 (0.177)
	Experience Sq	-0.027 (0.037)	-0.058 (0.043)	-0.072 (0.047)	-0.017 (0.053)	0.119* (0.065)	-0.138* (0.079)
	Marital Status	-0.058** (0.026)	-0.066** (0.026)	-0.049 (0.033)	-0.093** (0.039)	-0.051 (0.050)	-0.059 (0.054)
	No. of Family	0.075 (0.064)	0.099* (0.057)	0.055 (0.078)	0.254*** (0.095)	0.053 (0.122)	-0.132 (0.148)
	Log Migration Cost	0.069 (0.541)	0.931 (0.632)	0.280 (0.850)	-0.316 (0.847)	-0.327 (0.830)	-0.570 (0.894)
	_cons	-0.133 (0.547)	-0.989 (0.641)	-0.409 (0.851)	-0.012 (0.853)	0.308 (0.846)	0.553 (0.919)
N	*Group 1*				289		
	Group 2				137		

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

APPENDIX C
DISTRIBUTIONAL WAGE DECOMPOSITION (READING)

	DV- Log Wage	RIF Mean	RIF Q (10)	RIF Q (25)	RIF Q (50)/ Median	RIF Q (75)	RIF Q (90)
Overall	Group_1 Competent	7.783*** (0.059)	7.542***(0.070)	7.666*** (0.084)	7.804*** (0.087)	7.917***(0.088)	8.085***(0.117)
	Group_2 Non-Competent	7.636*** (0.011)	7.378***(0.014)	7.497*** (0.016)	7.675***(0.017)	7.850***(0.018)	7.949***(0.018)
	Difference	0.147** (0.060)	0.165** (0.071)	0.169** (0.085)	0.129 (0.089)	0.067 (0.090)	0.136 (0.118)
	Explained	0.061** (0.025)	0.045**(0.021)	0.079** (0.032)	0.086** (0.037)	0.049 (0.035)	0.052* (0.031)
	Unexplained	0.086 (0.060)	0.120* (0.070)	0.090 (0.086)	0.043 (0.091)	0.018 (0.093)	0.084 (0.120)
Explained	Education	0.031** (0.014)	0.017 (0.015)	0.040** (0.020)	0.032 (0.020)	0.018 (0.024)	0.051* (0.027)
	Experience	0.004 (0.031)	0.003 (0.019)	0.005 (0.037)	0.007 (0.047)	0.004 (0.032)	0.002 (0.015)
	Experience Sq	0.022** (0.011)	0.018* (0.010)	0.030** (0.015)	0.039** (0.019)	0.020* (0.011)	0.001 (0.009)
	Marital Status	-0.003 (0.004)	0.007 (0.007)	-0.001 (0.006)	-0.002 (0.006)	-0.006 (0.007)	-0.002 (0.006)
	No. of Family	0.005 (0.006)	0.002 (0.003)	0.003 (0.004)	0.006 (0.007)	0.010 (0.011)	0.001 (0.004)
	Log Migration Cost	0.001 (0.003)	-0.003 (0.003)	0.002 (0.005)	0.003 (0.005)	0.003 (0.005)	-0.001 (0.005)
Unexplained	Education	-0.057 (0.197)	0.119 (0.199)	-0.016 (0.250)	-0.131 (0.265)	-0.514 (0.327)	0.105 (0.544)
	Experience	-0.841 (0.658)	-0.184 (0.266)	-1.259 (0.871)	-1.871* (1.086)	-1.114 (0.956)	0.735 (1.053)
	Experience Sq	0.556 (0.488)	0.109 (0.174)	0.813 (0.676)	1.209 (0.741)	0.773 (0.690)	-0.392 (0.709)
	Marital Status	0.039 (0.089)	0.074 (0.120)	0.057 (0.156)	-0.105 (0.130)	-0.041 (0.146)	0.170 (0.148)
	No. of Family	-0.129 (0.372)	-0.101 (0.156)	-0.075 (0.355)	-0.059 (0.551)	-0.040 (0.545)	0.030 (0.735)
	Log Migration Cost	-1.258 (2.875)	1.363 (2.158)	0.249 (5.003)	1.008 (3.995)	-0.053 (4.567)	-3.656 (4.267)
	_cons	1.776 (2.847)	-1.260 (2.287)	0.319 (4.993)	-0.007 (4.041)	1.008 (4.438)	3.093 (4.119)
N	*Group 1*				18		
	Group 2				408		

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

APPENDIX D
DISTRIBUTIONAL WAGE DECOMPOSITION (WRITING)

	DV- Log Wage	RIF Mean	RIF Q (25)	RIF Q (50)/ Median	RIF Q (75)	RIF Q (90)
Overall	Group_1 Competent	7.793*** (0.083)	7.668*** (0.107)	7.873*** (0.129)	7.913*** (0.133)	8.083*** (0.135)
	Group_2 Non-Competent	7.639*** (0.011)	7.500*** (0.016)	7.677*** (0.017)	7.854*** (0.018)	7.952*** (0.017)
	Difference	0.154* (0.084)	0.168 (0.108)	0.195 (0.130)	0.059 (0.134)	0.131 (0.136)
	Explained	0.040 (0.037)	0.040 (0.043)	0.048 (0.056)	0.038 (0.049)	0.030 (0.032)
	Unexplained	0.114 (0.094)	0.128 (0.127)	0.148 (0.154)	0.021 (0.138)	0.102 (0.140)
Explained	Education	0.034** (0.015)	0.039** (0.019)	0.034* (0.020)	0.028 (0.023)	0.050* (0.027)
	Experience	-0.049 (0.048)	-0.058 (0.057)	-0.074 (0.072)	-0.051 (0.050)	-0.024 (0.025)
	Experience Sq	0.038*** (0.014)	0.049*** (0.019)	0.065*** (0.024)	0.033** (0.015)	0.003 (0.015)
	Marital Status	-0.004 (0.006)	-0.001 (0.009)	-0.003 (0.010)	-0.010 (0.011)	-0.001 (0.010)
	No. of Family	0.020** (0.009)	0.009 (0.010)	0.025** (0.012)	0.036** (0.016)	0.002 (0.014)
	Log Migration Cost	0.000 (0.001)	0.001 (0.003)	0.001 (0.002)	0.001 (0.003)	-0.001 (0.002)
Unexplained	Education	-0.947** (0.420)	-0.396 (0.363)	-0.552 (0.695)	-1.737** (0.676)	-2.192*** (0.491)
	Experience	0.595 (0.449)	1.812*** (0.494)	0.931 (0.747)	0.558 (0.702)	1.059** (0.520)
	Experience Sq	-0.869** (0.350)	-1.825*** (0.485)	-1.324** (0.569)	-0.944* (0.509)	-1.463*** (0.462)
	Marital Status	0.417** (0.193)	0.547*** (0.187)	0.447 (0.309)	0.350 (0.291)	0.729*** (0.240)
	No. of Family	0.306 (0.647)	-0.702 (0.552)	0.318 (1.069)	0.508 (1.023)	1.891*** (0.728)
	Log Migration Cost	0.258 (5.314)	2.626 (4.611)	1.702 (8.875)	0.920 (8.471)	2.729 (5.905)
	_cons	0.354 (5.334)	-1.935(4.595)	-1.376 (8.870)	0.366 (8.470)	-2.651 (5.928)
N	*Group 1*	9				
	Group 2	417				

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

REFERENCES

Autor DH, Handel MJ (2013) Putting tasks to the test: human capital, job tasks, and wages. J Labor Econ 31(S1):S59–S96. https://doi.org/10.1086/669332
Becker GS (1964) Human capital: a theoretical and empirical analysis, with special reference to education. University of Chicago Press, Chicago
Bleakley H, Chin A (2004) Language skills and earnings: evidence from childhood immigrants. Rev Econ Stat 86(2):481–496. https://doi.org/10.1162/003465304323031094
Borjas GJ (2003) The labour demand curve is downward sloping: reexamining the impact of immigration on the labour market. Q J Econ 118(4):1335–1374. https://doi.org/10.1162/003355303322552810
Chiswick BR, Miller PW (1995) The endogeneity between language and earnings: international analyses. J Labor Econ 13(2):246–288. https://doi.org/10.1086/298406
Council of Europe (2001) Common European framework of reference for languages: learning, teaching, assessment. Cambridge University Press, Cambridge
Devadason ES (2021) Foreign labour policy and employment in manufacturing: the case of Malaysia. J Contemp Asia 51(3):389–412. https://doi.org/10.1080/00472336.2020.1730680
Doeringer PB, Piore MJ (1971) Internal labor markets and manpower analysis. Heath Lexington Books, Lexington
Dustmann C (1994) Speaking fluency, writing fluency and earnings of migrants. J Popul Econ 7(2):133–156. https://doi.org/10.1007/BF00173615
Dustmann C, Fabbri F (2003) Language proficiency and labour market performance of immigrants in the UK. Econ J 113(489):695–717. https://doi.org/10.1111/1468-0297.00151
Dustmann C, Van Soest A (2002) Language and the earnings of immigrants. Ind Labor Relat Rev 55(3):473–492. https://doi.org/10.1177/001979390205500304
Dustmann C, Frattini T, Preston IP (2013) The effect of immigration along the distribution of wages. Rev Econ Stud 80(1):145–173. https://doi.org/10.1093/restud/rds019
Firpo S, Fortin NM, Lemieux T (2009) Unconditional quantile regressions. Econometrica 77(3):953–973. https://doi.org/10.3982/ECTA6822
Liwiński J (2019) The wage effect of foreign language skills: evidence from Poland. Empirica 46(4):691–711. https://doi.org/10.1007/s10663-018-9417-2
Mincer J (1974) Schooling, experience, and earnings. Columbia University Press, New York
Piore MJ (1979) Birds of passage: migrant labor and industrial societies. Cambridge University Press, Cambridge
Rios-Avila F (2020) Recentered influence functions in Stata: methods for distributional analysis. Stata J 20(1):51–94. https://doi.org/10.1177/1536867X20909690