

Analysis and Forecasting of Youth Employment: A Comparative Study between Shandong Province, China and Klang Valley, Malaysia

Qingyu Kong¹, Seng Yue Wong^{2*}, Ran Li³

¹ Institute for Advanced Studies, Universiti Malaya, Kuala Lumpur, 50603, Malaysia

²Dean's Office, Academy of Malay Studies, Universiti Malaya, 50603 Kuala Lumpur, Malaysia

³Institute of China Studies, Faculty of Arts and Social Sciences, Universiti Malaya, 50603 Kuala Lumpur, Malaysia

Qingyu Kong1: 25054061@siswa.um.edu.my

Seng Yue Wong2*: wongsengyue@um.edu.my

Ran Li3: liran@um.edu.my

*Corresponding e-mail: wongsengyue@um.edu.my

HOW TO CITE:

Qingyu Kong, Seng Yue Wong,
Ran Li (2026). Analysis and
Forecasting of Youth Employment:
A Comparative Study between
Shandong Province, China and
Klang Valley, Malaysia
International Journal of Special
Education, 41(25s), 947 - 967

COPYRIGHT STATEMENT:

Copyright: © 2026 Authors.
Open access publication under the
terms and conditions
of the Creative Commons Attribution
(CC BY)
license (<http://creativecommons.org/licenses/by/4.0/>).

ABSTRACT:

Youth labor markets in Shandong Province, China, and the Klang Valley, Malaysia, have moved from problems of job quantity to problems of structural friction, yet the two regions are usually studied with incompatible indicators and in separate literatures. This article sets out a quantitative research design for analyzing and forecasting youth employment in both regions, with NEET (not in employment, education or training) status treated as an endogenous determinant of later outcomes rather than only as a residual outcome of weak demand. Four research questions are addressed. The first establishes harmonized baselines and trends since 2014 for labor force participation, unemployment and NEET prevalence among young people aged 15 to 29, together with survey-based measures of long-term disengagement. The second asks whether the duration of disengagement is associated with psychological distress, measured with the Kessler 6 scale, and with downgraded career expectations, measured as a reservation wage gap. The third uses multi-group structural equation modeling, preceded by measurement invariance testing, to compare how perceived involution and skill mismatch shape disengagement trajectories in each region. The fourth evaluates whether adding structural employability variables, including an education and industry matching index and exposure to digital substitution, improves the explanatory power and out-of-sample accuracy of NEET forecasting models relative to a macroeconomic baseline. The design combines official and survey-based secondary series with a harmonized survey of 1,500 school leavers aged 18 to 29, 750 in each region. Decision rules, power calculations, endogeneity checks and ethical safeguards are fixed before data collection. Because the two regions publish very different youth statistics, the article also shows how a comparison can be built when official series are missing or not comparable, and it sets out how the resulting models can be used for ex-ante policy scenarios.

Keywords: youth employment; NEET; scarring effect; involution; skill mismatch; multi-group structural equation modeling; forecasting; Shandong; Klang Valley

Introduction

Youth employment remains one of the most persistent policy problems in emerging Asian economies. Globally, the youth unemployment rate fell to 13 percent in 2023, its lowest level in fifteen years, but one in five young people aged 15 to 24 was still not in employment, education or training (International Labour Organization [ILO], 2024). Headline improvement thus coexists with a large population of young people who are outside both the labor market and the education system. In China, the urban surveyed unemployment rate for those aged 16 to 24, excluding students, stood at 18.9 percent in August 2026 (National Bureau of Statistics of China [NBS], 2026). This series has been published only since the method was revised to exclude students, beginning with December 2023 data, and it is available only at the national level (NBS, 2024a). In Malaysia, the unemployment rate for those aged 15 to 24 was 10.3 percent in 2024 and 10.1 percent in September 2025, against an overall rate of about 3 percent (Department of Statistics Malaysia [DOSM], 2025a, 2025b). In Selangor, the state that holds most of the Klang Valley population, overall unemployment was only 1.8 percent in the third quarter of 2025 (DOSM, 2025c). Low headline rates in the country's most developed region can sit alongside persistent difficulties for the young.

These national figures conceal regional mechanisms that differ in kind. Shandong is one of China's largest producers of higher education graduates and a traditional industrial province in the middle of restructuring. Rapid tertiary expansion has outpaced the upgrading of local industrial demand, weakening the signaling value of degrees and channeling graduates into prolonged credential competition, including repeated attempts at postgraduate entrance and civil service examinations (Mok & Qian, 2018; Zhang et al., 2024). The everyday term for this pressure, *involution* (*neijuan*), describes escalating competitive investment that yields diminishing returns (Liu et al., 2022). Some young people respond by withdrawing into what

is effectively hidden NEET status, supported by their families while they prepare for examinations (B. H. Y. Chen et al., 2026; Lan et al., 2026). The Klang Valley, Malaysia's main urban and service economy with about 8.1 million residents (DOSM, n.d., 2022), shows a different pattern. Here the problem is less about withdrawal than about weak absorption: graduates take jobs below their qualifications, entry wages remain low relative to living costs, and platform-based gig work absorbs young people without offering protection or progression (Abu Rahim & Sivashanmugam, 2024; Siew & Sze, 2025; Uchiyama et al., 2022).

Comparative research on these two settings faces three problems. First, the indicators are not harmonized. Chinese and Malaysian statistics define youth age bands, student status and unemployment differently, and neither country publishes a regular youth NEET series for the regions studied here, so direct comparison of published rates can mislead. Second, NEET status is usually modeled as an outcome of macroeconomic conditions, even though a growing body of work shows that early disengagement leaves lasting economic and psychological scars that shape later trajectories (Caniglia, 2024; Filomena et al., 2025; Ralston et al., 2022). Third, forecasting models for youth labor markets rely mainly on lagging macroeconomic predictors such as output growth and inflation, and they rarely include structural variables that capture how well education supply matches industrial demand or how exposed entry-level jobs are to digital substitution (ILO, 2024; Tleppayev & Zeinolla, 2025).

This article responds with a research design for a comparative study of youth employment in Shandong and the Klang Valley. NEET status, and above all the duration of disengagement, sits at the center of the analysis. It is guided by four research questions:

RQ1: How can specific statistical measures (labor force participation rate, NEET prevalence and longitudinal duration of disengagement)

characterize the current baseline status and cross-temporal evolution of youth employment in Shandong Province and the Klang Valley?

RQ2: To what extent does NEET status, operationalized as an endogenous variable centered on the duration of disengagement, predict the degree of scarring in young people, specifically their psychological distress (Kessler 6 scale) and the downgrading of their career expectations (reservation wage gap)?

RQ3: Within a unified structural equation modeling framework, what statistically significant differences exist in the path coefficients of the core driving variables, perceived involution in Shandong and skill mismatch in the Klang Valley, with respect to their effects on youth employment trajectories?

RQ4: After structural employability variables (such as the digital skill substitution rate and an education and industry matching index) are integrated into forecasting models, how can gains in explanatory power (R^2) and reductions in prediction error provide a quantified basis for ex-ante policy intervention?

The corresponding objectives are to measure and compare baseline distributions and temporal dynamics of youth participation and NEET scale in the two regions (RO1); to estimate the strength of the association between NEET duration and the depletion of psychological resources and the revision of career expectations, using design features that strengthen causal interpretation (RO2); to identify universal pathways and regional differences in employment dynamics using multi-group structural equation modeling (RO3); and to test whether structural employability factors improve the accuracy of NEET forecasts and to use the improved models for ex-ante policy scenario simulation (RO4).

The article contributes in three ways. A harmonized indicator set allows two youth labor markets that are rarely studied together to be compared on the same terms. In the structural model, the driver attributed to each region is measured in both, so that path differences can be

tested instead of assumed. The forecasting strategy judges structural predictors by their out-of-sample accuracy, since that is what matters when forecasts are used for policy. Section 2 reviews the literature, Section 3 states the hypotheses, Section 4 describes the methods, and Section 5 discusses contributions and limitations.

2. Literature Review and Theoretical Framework

2.1 Youth Transition Regimes and the NEET Indicator

Walther (2006) argued that young people's movement from school to work is organized by institutional regimes that combine education, training, labor market regulation and welfare. Early applications of this framework relied on the youth unemployment rate. That indicator misses young people who have stopped searching or never entered the labor force. The NEET indicator counts them, since it includes inactive as well as unemployed youth who are outside education and training, and a large literature now links NEET status to family background, educational trajectories, health and local labor demand (Rahmani & Groot, 2023). Comparative work in Europe has shown that NEET rates respond to configurations of policy: similar macroeconomic conditions produce different levels of youth vulnerability depending on institutional arrangements (Lauri et al., 2025).

Transition regime typologies were developed for European welfare states, and recent reviews note that most NEET research still comes from high-income Western countries (Gunnes et al., 2025). Two features of the Chinese and Malaysian settings sit uneasily with these typologies. In China, family support and a strong examination culture allow disengagement to take the form of extended preparation for selective credentials, which official statistics may not record as unemployment. In Malaysia, a large informal and platform sector allows young people to remain technically employed while accumulating little human capital. A comparative design needs indicators that capture both hidden withdrawal

and low-quality participation, and it needs a common analytical structure in which these different forms of disengagement can be compared.

2.2 Scarring Effects of Early Disengagement

The scarring hypothesis holds that early periods of joblessness have lasting consequences beyond their immediate duration. Using German administrative data, Schmillen and Umkehrer (2017) showed that early-career unemployment predicts more unemployment later in working life. Filomena et al. (2025) reported that early nonemployment is associated with weaker labor market outcomes in adulthood, and Ralston et al. (2022) found that NEET status in Scotland marks long-term disadvantage that persists after other characteristics are taken into account. Fabrizi and Rocca (2024) showed that the duration of NEET spells, and not only their occurrence, depends on socioeconomic background. For this reason the present design measures duration rather than a binary NEET indicator.

Scarring also has a psychological dimension. A meta-analysis found that NEET status is associated with mental ill-health and psychological distress, although longitudinal studies more often show poor mental health preceding NEET status than the reverse (Gariépy et al., 2022). Using United States cohort data, Caniglia (2024) found that young adults who were NEET in their late teens and twenties reported more depressive symptoms into early midlife, a result that held after propensity score matching. Wright et al. (2021) found that the association between youth unemployment and later mental health is strongest among people at the poorer end of the mental health distribution. Resource-based accounts suggest that NEET status depletes material, personal and social resources together (X. Chen & Chan, 2025). A further mechanism concerns aspirations. Mühlböck et al. (2022) found that young people adjust their vocational aspirations and expectations during unemployment, and research on reservation wages shows that the lowest wage a job seeker

will accept declines at a modest rate as a spell lengthens (Krueger & Mueller, 2016). Tomlinson and Tholen (2023) argue that these individual adjustments interact with employer behavior, since firms raise entry thresholds in difficult periods. Scarring is treated here as two related outcomes, psychological distress and the downgrading of career expectations, and it treats the direction of the distress association as an empirical question to be handled in the analysis.

2.3 Involution and Credential Competition in Shandong

Massification of higher education in China has increased graduate supply faster than the growth of graduate-level jobs, producing skills mismatch, informal employment among graduates and credential inflation (Mok & Qian, 2018). Zhang et al. (2024) describe a signaling paradox in which overeducation lowers the returns to degrees while increasing the pressure to acquire further credentials. Lan et al. (2026) find that the link between education and job quality among Chinese youth has weakened, which they frame as a meritocracy trap. Involution is the subjective counterpart of these structural conditions. Liu et al. (2022) developed and validated a competition psychology scale that measures it among university students, network analysis suggests that perceived involution is closely tied to lower subjective well-being (Li, 2025), and a recent survey of undergraduates links academic involution to how students perceive their own employability (Ji & Li, 2026). In interviews, young Chinese graduates expressed a strong preference for stable jobs, and most described trying to balance intense competition against lying flat (Deng et al., 2025). Withdrawal from competition has been linked to frustrated needs for autonomy (Lu et al., 2025) and to anxiety and depression (Ren et al., 2026), and pursuit of further study, including abroad, can be understood partly as a defensive response to involution (B. H. Y. Chen et al., 2026).

If involution works in this way, it should lengthen disengagement. Young people who expect additional effort to bring no advantage may delay

labor market entry, extend examination preparation or stop searching. The pathway should be strongest in Shandong, where public sector jobs carry high status and attract intense competition.

2.4 Skill Mismatch and Gig Absorption in the Klang Valley

Malaysian research has documented persistent graduate underemployment. Khazanah Research Institute analysis points to widespread overqualification among fresh graduates (Abu Rahim & Sivashanmugam, 2024) and to slow career progression among skilled workers (Hamid et al., 2024). Using national data for 2020 to 2023, Siew and Sze (2025) estimate that about one third of Malaysian graduates were underemployed, mostly in service and sales occupations, and survey evidence suggests that headline graduate employability rates hide substantial overeducation (Low & Mah, 2024). In the Klang Valley, high living costs and low entry wages make platform work an accessible option, but such work lacks social protection and transfers risk to workers (Uchiyama et al., 2022), and recent conflicts between delivery riders and platforms show how weak the bargaining position of these workers remains (Uchiyama et al., 2025). Mismatch can lengthen disengagement indirectly: young people who cannot find work that uses their training may cycle between short gig spells and inactivity, or give up searching for graduate-level positions.

Measuring skill mismatch requires care. Bischof (2024) compared several approaches and found that subjective self-assessments and objective comparisons between qualifications and job requirements capture different aspects of mismatch. The present design combines a subjective scale with an objective indicator based on the match between education level and the skill level of the most recent occupation.

2.5 Forecasting Youth Labor Market Outcomes

Most youth labor market projections, including international ones, extrapolate from aggregate

demand indicators (ILO, 2024). Such models are informative about cyclical movements but less so about structural change. Panel evidence from eight Central Asian and Caucasus economies links tertiary enrolment and youth population growth to youth unemployment (Tleppayev & Zeinolla, 2025), which suggests that education-side variables belong in forecasting models. Two such variables are used here. The first is the alignment between the fields in which graduates are trained and the sectors in which employment is growing. The second is the exposure of entry-level occupations to automation and artificial intelligence, which can be measured with occupation-level exposure indices (Eloundou et al., 2024; Felten et al., 2021). If these variables carry information beyond macroeconomic conditions, adding them should improve forecasts. Whether they do has to be judged out of sample, because in-sample R^2 rises whenever predictors are added (Hyndman & Athanasopoulos, 2021).

2.6 Integrated Framework

The study draws on two theoretical perspectives. Ecological systems theory (Bronfenbrenner, 1979) treats development as the product of nested systems. Here the macro level comprises labor market structure and institutions, the meso level comprises field-level pressures such as credential devaluation and gig platform expansion, and the micro level comprises individual perceptions, disengagement and psychological responses. The capability approach (Sen, 1999) reframes NEET status as a restriction of the real opportunities available to young people rather than as an individual deficit. Combined, the two perspectives imply a sequence: structural pressures are experienced as perceived constraints, perceived constraints lengthen disengagement, and long disengagement goes with depleted psychological resources and lower expectations. Ecological systems theory supplies the levels and the capability approach supplies the interpretation of NEET status. A human capital framework alone was considered and set aside, because it treats disengagement as

depreciation of skills and has little to say about

how young people perceive their options.

Figure 1 presents the resulting model. Perceived involution and skill mismatch are measured in both regions and are expected to predict the duration of disengagement. Duration is expected to predict psychological distress and the reservation wage gap, and distress is expected to carry part of the association between duration and expectations. The model is estimated for Shandong and the Klang Valley within a multi-group framework, which allows the same paths to be compared directly. Region-level structural employability variables enter the forecasting component.

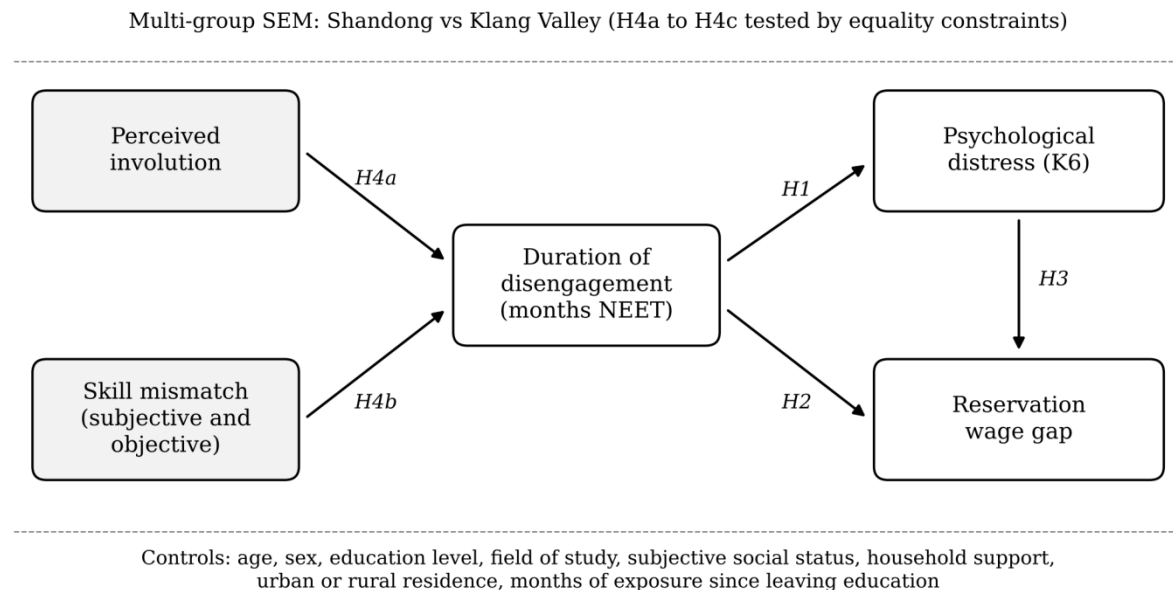


Figure 1. Conceptual model linking structural drivers, duration of disengagement and scarring outcomes, estimated in parallel for Shandong and the Klang Valley.

3. Hypotheses

RQ1 is descriptive and has no directional hypothesis. The hypotheses below address RQ2 to RQ4. H4a and H4b are directional comparisons and are supported only if the regional difference is significant and runs in the predicted direction.

H1. Longer duration of disengagement is associated with higher psychological distress (K6), net of controls, in both regions.

H2. Longer duration of disengagement is associated with a larger reservation wage gap, indicating downgraded career expectations, in both regions.

H3. Psychological distress partially mediates the association between duration of disengagement

and the reservation wage gap.

H4a. Perceived involution is positively associated with duration of disengagement, and this path is stronger in Shandong than in the Klang Valley.

H4b. Skill mismatch is positively associated with duration of disengagement, and this path is stronger in the Klang Valley than in Shandong.

H4c. The indirect associations of perceived involution and skill mismatch with psychological distress and the reservation wage gap, through duration of disengagement, differ between the two regions.

H5. Forecasting models that add structural employability variables to a macroeconomic baseline produce lower out-of-sample errors for youth NEET rates than the baseline models, in

the Klang Valley and in Shandong.

4. Method

4.1 Design Overview

The design is comparative and quantitative, with two data streams and a qualitative follow-up. Stream A consists of official and survey-based secondary statistics from 2014 onward and serves RQ1 and RQ4. Stream B is a harmonized cross-sectional survey with a retrospective activity calendar and serves RQ1, RQ2 and RQ3. A two-

wave panel would have given stronger evidence on the order of disengagement and distress, but it would have doubled fieldwork time and exposed a hidden population to heavy attrition. The cross-sectional design with a 36-month calendar was chosen instead, at the cost of relying on recall. Interviews with participants selected from Stream B are planned afterward to interpret patterns that the models leave unexplained. Table 1 maps each research question to its objective, hypotheses, data and analysis.

Table 1. Alignment of Research Questions, Objectives, Hypotheses, Data and Analysis

RQ	RO	Hypotheses	Data	Analysis
RQ1	RO1	Descriptive	Stream A (Klang Valley annual 2014 to 2025; Shandong biennial 2014 to 2022); Stream B weighted estimates	Harmonized indicators; trend and break analysis; survey-based duration distributions
RQ2	RO2	H1, H2, H3	Stream B (650 usable cases per region)	Structural equation model with bootstrapped indirect effects; propensity weighting and sensitivity analysis
RQ3	RO3	H4a, H4b, H4c	Stream B (650 usable cases per region)	Measurement invariance testing; multi-group SEM with equality constraints
RQ4	RO4	H5	Stream A panels (192 state-years; 125 province-waves), checked against Stream B	Nested panel forecasting models; rolling-origin evaluation; scenario simulation to 2030

Note. Stream A = official and survey-based secondary statistics; Stream B = harmonized regional survey. The Malaysian panel covers 16 states and federal territories from 2014 to 2025; the Chinese panel covers the 25 provinces sampled by the China Family Panel Studies in five waves from 2014 to 2022.

4.2 Study Settings and Definitions

Shandong. Shandong Province is defined by its administrative boundary and covers 16 prefecture-level cities.

Klang Valley. The Klang Valley has no single statutory boundary, and studies define it in different ways. This study follows the definition of Greater Kuala Lumpur and the Klang Valley adopted in Malaysia's Economic Transformation Programme, which identifies the area through 10 local authorities: Kuala Lumpur City Hall, Putrajaya Corporation, and the Shah Alam,

Petaling Jaya, Subang Jaya, Klang, Kajang, Selayang, Ampang Jaya and Sepang councils (Performance Management and Delivery Unit, 2010; Yasin et al., 2022). Because survey screening and census data are organized by administrative district rather than by local authority, the 10 authorities are mapped to seven units: the Federal Territories of Kuala Lumpur and Putrajaya and the Selangor districts of Petaling, Klang, Gombak, Hulu Langat and Sepang. The Selangor districts of Kuala Langat, Kuala Selangor, Hulu Selangor and Sabak Bernam, which lie outside the 10 authorities, are

excluded. A wider definition that adds some of these districts is also in use, but they are largely semi-rural, and their labor markets differ from the metropolitan core that defines the Klang Valley as an employment region. Table 2 lists the seven units, their 2020 census populations and the survey allocation derived from them. The

same boundary is applied wherever the data allow. Where official statistics are available only for Selangor as a whole, Kuala Lumpur, Putrajaya and Selangor figures are combined with population weights, and this approximation is reported.

Table 2. Study Boundary of the Klang Valley and Survey Allocation

Administrative unit	Local authorities within the Economic Transformation Programme boundary	Population, 2020 census	Share (%)	Survey allocation (n)
W.P. Kuala Lumpur	Kuala Lumpur City Hall	1,982,112	24.3	182
W.P. Putrajaya	Putrajaya Corporation	109,202	1.3	10
Petaling district	Petaling Jaya, Subang Jaya and most of Shah Alam	2,298,130	28.2	212
Klang district	Klang and the remainder of Shah Alam	1,088,942	13.4	100
Gombak district	Selayang and part of Ampang Jaya	942,400	11.6	87
Hulu Langat district	Kajang and part of Ampang Jaya	1,400,461	17.2	129
Sepang district	Sepang	325,244	4.0	30
Total	10 local authorities	8,146,491	100.0	750

Note. Population figures are from the 2020 Population and Housing Census of Malaysia (DOSM, n.d., 2022); the figure for W.P. Putrajaya is that of the single parliamentary constituency P.125, which covers the whole territory. Allocation is proportional to population share and sums to the 750 responses to be collected in the Klang Valley. Units are not analyzed separately.

Youth and NEET. Youth is defined as ages 15 to 29, with separate reporting for ages 15 to 24 (the international convention) and 25 to 29 (early career). The upper band is included because extended education and delayed entry are common in both settings, and a 15 to 24 definition would miss most graduates who are preparing for examinations. The secondary analyses in Stream A use the full 15 to 29 range. The survey in Stream B covers ages 18 to 29, and where survey and secondary indicators overlap

they are compared for the same ages. Where an official series uses a different lower bound, as the Chinese national series does by starting at 16, the difference is noted, not adjusted. A young person is classified as NEET if they are not enrolled in formal education, not employed, and not participating in structured training. Two extensions are reported separately and kept out of the core definition, which keeps results comparable with official statistics: young people preparing full time for examinations without

enrollment (hidden withdrawal), and young people whose only work is irregular platform or gig work without social protection (marginal participation).

4.3 Stream A: Secondary Indicators

The two regions differ in the youth statistics that are available, and the design adapts to the data that exist. For the Klang Valley, the annual Labour Force Survey reports and quarterly principal statistics give participation and unemployment by state, but the published report does not present state-by-age tables (DOSM, 2025b). Youth indicators for all states and federal territories, including Kuala Lumpur, Putrajaya and Selangor, are therefore computed for 2014 to 2025 from Labour Force Survey microdata requested from DOSM. For Shandong, NBS publishes youth surveyed unemployment only at the national level (NBS, 2024a), and the

provincial table of registered urban unemployment is no longer released (NBS, 2024b). The census and the 1 percent sample survey give provincial detail only for 2015 and 2020, which is too few points for a trend. Shandong youth indicators are therefore estimated from the China Family Panel Studies, a national longitudinal survey with provincial identifiers (Xie & Hu, 2014), for the waves of 2014, 2016, 2018, 2020 and 2022. Because provincial subsamples are small, these direct estimates are combined with provincial auxiliary data from the statistical yearbooks, the 2015 1 percent sample survey and the 2020 census in an area-level small area estimation model (Fay & Herriot, 1979; Hadam et al., 2024). ILOSTAT series are used as external benchmarks. Each indicator is harmonized according to the rules in Table 3, and all rules and transformations are documented in a public codebook.

Table 3. Harmonized Indicator Set for RQ1

Indicator	Definition	Shandong source	Klang source	Valley	Harmonization issue
Labor force participation rate	Labor force as share of youth population, ages 15 to 24 and 25 to 29	CFPS waves 2014 to 2022 with small area estimation	Labour Force Survey microdata, 2014 to 2025	Force	Biennial in Shandong, annual in the Klang Valley; compared at common years
Youth unemployment rate	Unemployed as share of youth labor force	CFPS-based provincial estimates; national NBS series (ages 16 to 24) as context	Labour Force Survey microdata	Force	National Chinese series excludes students from December 2023 onward and is not used as a provincial value
NEET rate	Share of youth not in education, employment or training	CFPS-based provincial estimates	Labour Force Survey microdata	Force	Training definitions differ; sensitivity check with narrow and broad definitions; Stream B gives a rate among school leavers only

Indicator	Definition	Shandong source	Klang source	Valley	Harmonization issue
Long-term disengagement share	Share of NEET youth disengaged 12 months or longer	Stream B calendar	Stream B calendar		Not published officially; survey-based estimate, ages 18 to 29, with confidence interval
Graduate mismatch rate	Share of employed tertiary graduates in jobs below their skill level	Stream B; CFPS and education codes	Stream B; Labour Force occupation and education codes	ISCED and ISCO mapping and identically	

Note. CFPS = China Family Panel Studies; NBS = National Bureau of Statistics of China.

Trends are described from 2014 onward, annually for the Klang Valley to 2025 and biennially for Shandong to 2022. For the Klang Valley, interrupted time-series models with a break in 2020 test whether the pandemic shifted the level or slope of each series. For Shandong, the 2014 to 2018 waves are compared with the 2020 and 2022 waves. Regional differences in trend are tested at the common years by pooling the series with region-by-period interactions. Because the series are short, these analyses are descriptive and are reported with confidence intervals. They are not forecasts.

4.4 Stream B: Harmonized Survey

Population and eligibility. The survey targets young people aged 18 to 29 who have completed at least upper secondary education, who left full-time education at least six months before the survey, and who have lived in Shandong or in one of the seven Klang Valley units in Table 2 for at least the previous 12 months. Current full-time students are excluded, because the scarring outcomes refer to the period after leaving education. The lower age limit of 18 avoids the need for parental consent and covers the period after compulsory schooling, when disengagement is most consequential. Residence is verified through screening questions on city or district and postcode, and respondents living outside the defined boundary are screened out.

Sampling. Because NEET youth are a dispersed

and partly hidden population, probability sampling is not feasible within the study's resources. The survey uses quota sampling through commercial online panels in each region, supplemented by targeted recruitment through youth employment services, university alumni networks and community groups. In the Klang Valley, the 750 responses are allocated across the seven units as shown in Table 2. In Shandong, the 750 responses are allocated across the 16 prefecture-level cities in proportion to their resident population in the 2020 census. Within each region, quotas are also set by age band (18 to 24 and 25 to 29), sex and education level (tertiary or not), following census proportions. Activity status quotas ensure that at least 40 percent of the usable sample in each region, that is at least 260 of 650 respondents, is currently NEET or has been NEET for three months or longer during the previous 36 months, while the remainder provides a comparison group with no or brief disengagement. After data collection, raking weights align each sample with census distributions of age, sex and education among school leavers. Because the NEET group is deliberately oversampled, its weight is set to the regional NEET share estimated in Stream A. No official margin exists for either region. Weighted and unweighted estimates are both reported.

Sample size. The binding requirement is the ability to detect differences between regions in structural paths. For standardized coefficients,

the standard error of a between-group difference with 650 cases per group is approximately 0.055, which gives power of 0.80 at $\alpha = .05$ to detect a difference of about 0.16. The target is therefore 650 usable responses per region. Allowing for exclusions of about 15 percent for failed attention checks, duplicate responses and ineligibility, 750 responses per region, 1,500 in total, are to be collected. Before fieldwork, a Monte Carlo simulation based on the full model and plausible parameter values will confirm power for the key paths and for the indirect effects (Muthén & Muthén, 2002; Wang & Rhemtulla, 2021).

Instrument development. The questionnaire is prepared in English and translated into Simplified Chinese and Malay using forward translation, independent back-translation and expert review (Brislin, 1970). Shandong respondents complete the Chinese version. Klang Valley respondents may choose English, Malay or Simplified Chinese, reflecting the main languages of the area. Cognitive interviews with 10 young people in each region and a pilot with 50 respondents in each region will be used to check comprehension, completion time and preliminary reliability. Pilot respondents are not included in the main sample, and the main survey begins only when the core scales reach acceptable internal consistency in the pilot.

Data quality. Two attention-check items, a minimum completion time, device fingerprinting to remove duplicates, and consistency checks between the activity calendar and current status are applied. Missing data on scale items are handled by full information maximum likelihood under a missing at random assumption, with a sensitivity analysis using multiple imputation.

4.5 Measures

Table 4 summarizes the measures. Two design choices are central. First, both perceived involution and skill mismatch are measured in Shandong and in the Klang Valley. The research questions associate involution with Shandong and mismatch with the Klang Valley, but path coefficients can only be compared across groups if the same constructs are measured in each group. Measuring both drivers in both regions makes it possible to test whether involution matters more in Shandong and mismatch more in the Klang Valley. Second, duration of disengagement is constructed from a retrospective monthly activity calendar covering the previous 36 months, or the months since leaving full-time education if fewer. Months of exposure enter the model as a control, and the share of exposed months spent NEET is used in a robustness check, so that recent school leavers are not scored as having short spells only because they have had less time.

Table 4. Constructs and Measures

Construct	Measure	Items and scoring	Role
Perceived involution	Competition psychology scale for college students (Liu et al., 2022), adapted for non-student youth	12 items, 5-point agreement	Exogenous driver
Skill mismatch, subjective	Self-assessed underuse of skills and qualifications in current or most recent job	3 items, 5-point agreement	Exogenous driver

Construct	Measure	Items and scoring	Role
Skill mismatch, objective	Comparison of ISCED education level with ISCO skill level of current or most recent occupation (Bischof, 2024)	1 = education exceeds job requirement, 0 = otherwise	Exogenous driver (observed indicator)
Duration of disengagement	Months spent NEET in the previous 36 months or since leaving full-time education, from activity calendar; number of NEET spells	Continuous (0 to 36); count; months of exposure as control	Endogenous mediator
Psychological distress	Kessler 6 (Kessler et al., 2002)	6 items, 0 to 4, past 30 days; total 0 to 24	Outcome
Reservation wage gap	Log of lowest acceptable monthly wage when leaving education or last job minus log of current lowest acceptable wage, both deflated to 2025 prices	Continuous; positive values indicate downgraded expectations	Outcome
Occupational expectation	Expected occupation at exit and at present, coded to ISCO major group	Ordinal difference	Supplementary outcome
Controls	Age, sex, education level, field of study (STEM or not), subjective social status ladder (Adler et al., 2000), household financial support, urban or rural residence	Single items	Covariates

The K6 is short, captures nonspecific distress rather than a diagnosis, and has shown a stable factor structure and longitudinal invariance in a large Chinese sample (Zhang & Li, 2022). That sample consisted of older adults and the study found two factors, so a one-factor and a two-factor model are compared in each region, and invariance across the two youth samples is tested, not assumed. It is used as a screening scale for distress, not as a clinical instrument. The reservation wage gap follows the logic of reservation wage research (Krueger & Mueller, 2016). Because the starting value is reported retrospectively, the two wage values are anchored

to prevailing regional median starting salaries presented in the question, and recall bias is examined by comparing respondents whose exit was recent with those whose exit was earlier.

4.6 Analysis

Measurement models. Confirmatory factor analysis is estimated for each region using robust maximum likelihood. Covariance-based estimation is used rather than partial least squares, because the group comparisons depend on invariance tests and global fit indices that partial least squares does not provide in the same form. Model fit is judged with the comparative fit index ($CFI \geq .95$)

preferred, $\geq .90$ acceptable), the root mean square error of approximation ($RMSEA \leq .06$ preferred, $\leq .08$ acceptable) and the standardized root mean square residual ($SRMR \leq .08$) (Hu & Bentler, 1999). Composite reliability between .70 and .95, average variance extracted of at least .50 and heterotrait-monotrait ratios below .85 are required, and these criteria are reported with bootstrapped confidence intervals (Cheung et al., 2024). Because all survey measures come from one source, procedural remedies are built into the questionnaire, and common method variance is assessed with a marker variable and with a latent method factor (Podsakoff et al., 2024).

Measurement invariance. Before paths are compared, configural, metric and scalar invariance across the two regions are tested in sequence. A decrease in CFI of more than .010, together with an increase in RMSEA of more than .015, indicates non-invariance at a given level (F. F. Chen, 2007; Leitgöb et al., 2023). Metric invariance is the minimum condition for comparing structural paths, and scalar invariance is needed for comparing latent means. If full invariance fails, partial invariance is tested by freeing parameters identified through modification indices, provided that at least two loadings per factor remain constrained (Boer et al., 2018). The involution scale was developed with Chinese students and may not hold its structure among Klang Valley respondents. If metric invariance fails for that scale, the involution path cannot be compared across regions, and H4a is reported as untestable rather than as unsupported.

RQ1, survey component. For each region, weighted estimates are produced for the distribution of disengagement duration, the share of NEET respondents disengaged for 12 months or longer, and the shares of hidden withdrawal and marginal participation, all for school leavers aged 18 to 29 and with confidence intervals. Because current students are excluded, these figures describe school leavers and are not population NEET rates. They add duration and the two extended categories, which the secondary series in Table 3 cannot measure.

RQ2. A structural equation model regresses psychological distress and the reservation wage gap on duration of disengagement, with distress also predicting the wage gap, and with all controls included. Indirect effects are tested with 5,000 bias-corrected bootstrap samples (Hayes, 2022). Duration may be endogenous, because young people who are already distressed may remain disengaged longer, and meta-analytic evidence suggests that this direction is common (Gariépy et al., 2022). Three steps address this. First, the model controls for retrospectively reported distress and job search difficulty at the point of exit. Second, generalized propensity score weights for duration, estimated from pre-exit characteristics, are applied in a sensitivity model (Hirano & Imbens, 2004), and a matching estimator on a binary indicator of disengagement of 12 months or longer is used as a further check (Greifer & Stuart, 2021). Third, E-values are reported to show how strong an unmeasured confounder would need to be to explain away the observed association (VanderWeele & Ding, 2017). Because the survey is cross-sectional, an association that survives these checks is read as consistent with scarring, not as proof of it.

RQ3. The full model in Figure 1 is estimated as a multi-group structural equation model with region as the grouping variable. The baseline model allows all structural paths to differ between regions. Each hypothesized path is then constrained to equality in turn, and the change in model fit is evaluated with a scaled chi-square difference test. A significant deterioration indicates a regional difference in that path. A non-significant test is not read as evidence that the paths are equal. Where equality matters for interpretation, two one-sided tests against a smallest difference of interest of 0.10 in standardized units are reported. The same approach, applied to bootstrapped indirect effects, tests H4c. Because several paths are compared, the false discovery rate is controlled with the Benjamini and Hochberg (1995) procedure. Results are reported as unstandardized and standardized coefficients with confidence intervals for each region, so that

readers can judge the size of differences as well as their significance.

RQ4. The forecasting component models youth NEET rates and youth unemployment rates for ages 15 to 29. Regional series alone are too short for reliable forecasting with several predictors, so models are estimated on two panels and forecasts are then generated for the Klang Valley and Shandong. Each panel is fitted as a multilevel model with random intercepts for states or provinces and a common trend. Fixed effects with unit-specific trends were considered and rejected, because they would use most of the degrees of freedom in panels this short. The Malaysian panel covers the 16 states and federal territories annually from 2014 to 2025, giving 192

state-years. The Klang Valley forecast is the population-weighted combination of the forecasts for Kuala Lumpur, Putrajaya and Selangor. The Chinese panel covers the 25 provinces sampled by the China Family Panel Studies at the five waves from 2014 to 2022, giving 125 province-waves. Table 5 sets out the nested specifications. All predictors are lagged by one period, one year in Malaysia and one wave (two years) in China, to avoid using information that would not have been available at the time of the forecast. Graduate counts by field come from the education statistics of each country, allocated to states or provinces by the location of institutions, and occupational shares come from the Labour Force Survey and CFPS microdata.

Table 5. Nested Forecasting Specifications for RQ4

Model	Predictors	Purpose
M0	Autoregressive terms only	Naive benchmark
M1	M0 plus real GDP growth, CPI inflation, industry and services employment shares	Macroeconomic baseline
M2	M1 plus education and industry matching index	Contribution of supply and demand alignment
M3	M2 plus digital substitution exposure of entry-level employment and informal or platform employment share	Full structural model

Note. The matching index is computed as one minus the dissimilarity index between the distribution of new graduates across fields and the distribution of employment growth across corresponding sectors. Digital substitution exposure is the employment-weighted mean of an occupation-level artificial intelligence exposure score (Eloundou et al., 2024; Felten et al., 2021) for occupations typically entered by young workers.

Two kinds of evidence are reported. In-sample improvement is reported as the change in within-panel R^2 and adjusted R^2 from M1 to M3. Out-of-sample performance, which is the primary criterion for H5, is evaluated with rolling-origin cross-validation. In the Malaysian panel, models

are estimated on data up to year t and used to forecast years $t + 1$ to $t + 3$, with the origin moved forward one year at a time from 2019 to 2024 and horizons truncated at 2025, so that the first estimation window already contains five years. In the Chinese panel, models are estimated up to the 2018 and 2020 waves and used to forecast the next wave. Accuracy is summarized with root mean squared error, mean absolute error and mean absolute scaled error (Hyndman & Athanasopoulos, 2021). Differences in accuracy between M1 and M3 are tested with the Diebold and Mariano (1995) test with the small-sample correction of Harvey et al. (1997), applied to loss differentials pooled across units at each origin. If M3 does not outperform M1 out of

sample, H5 is rejected even if in-sample R^2 rises. Because the Chinese panel provides only two forecast origins, the Shandong evidence for H5 is treated as indicative and reported with this caveat.

The preferred model is then used to simulate NEET trajectories to 2030, annually for the Klang Valley from 2026 and at biennial intervals for Shandong from 2024, under three scenarios: a baseline that continues recent trends in the structural variables; an accelerated digital substitution scenario in which exposure rises by one standard deviation of its historical annual change; and a matching improvement scenario in which the matching index rises by one standard deviation of its historical annual change. Skill mismatch rates estimated from Stream B are used to check whether the structural variables are plausible for each region. Scenarios are reported with 80 and 95 percent prediction intervals. They are intended to show the direction and approximate size of the response to policy-relevant changes, not to provide point predictions.

4.7 Explanatory Qualitative Follow-up

After the quantitative analysis, 15 to 20 participants in each region will be invited to semi-structured interviews, and recruitment stops when no new themes emerge (Hennink & Kaiser, 2022). They are drawn from Stream B respondents who agreed to be contacted. Cases the model explains poorly are chosen on purpose: long disengagement with low distress, for example, or short disengagement with high distress. Each interview opens with a life-line drawn by the participant. The line anchors events in time, from leaving education to the present, and the transcripts are analyzed with reflexive thematic analysis (Braun & Clarke, 2022). Results are integrated with the quantitative findings at the levels of design, methods and interpretation, using joint displays that place path estimates alongside themes (Guetterman & Manojlovich, 2024).

4.8 Ethical Considerations

The study involves young adults, some of whom may be experiencing distress, and it asks about mental health and financial circumstances. Ethical approval will be obtained from the UMREC before any data are collected, and local requirements in both countries will be followed. Participants receive information in their chosen language and give informed consent online before starting the survey. No names or identification numbers are collected. An anonymous code lets each respondent withdraw their data before analysis. At the end of the survey, every respondent is shown information about free support services in their region, so that no one is singled out by their answers. Respondents who score 13 or above on the K6, the cut point proposed for serious psychological distress (Kessler et al., 2003), also see a short message encouraging them to contact these services. Interview participants may skip questions or stop at any time, and interviewers are trained to pause and provide referral information if distress becomes apparent. Data are stored on encrypted university servers, access is limited to the research team, and data will be deleted five years after the end of the project.

5. Discussion

5.1 Anticipated Contributions

The design responds to the problems identified in the introduction in four ways. First, the harmonized indicator set in Table 3 makes it possible to compare youth labor markets in Shandong and the Klang Valley on equivalent terms, and to show where official statistics are not comparable or not available. Hidden withdrawal and marginal participation are reported separately because they are the forms of disengagement that most clearly distinguish the two regions, and headline unemployment rates miss both.

Second, the model treats duration of disengagement as a mechanism rather than only as an outcome. If H1 to H3 are supported, the results would extend evidence on scarring, which

comes mainly from European and North American longitudinal data (Caniglia, 2024; Ralston et al., 2022; Schmillen & Umkehrer, 2017), to two Asian settings, and would show that scarring is associated with expectations as well as with mental health. Support for H3 would indicate that distress is one of the routes through which long disengagement relates to lower expectations, in which case psychological support would need to accompany employment services from the start rather than follow them. The reverse direction remains possible. If young people who are already distressed lower their expectations and stay out longer, the same pattern would appear, and the sensitivity analyses in Section 4.6 can narrow this explanation but not exclude it.

Third, measuring both drivers in both regions changes the comparative question. Earlier work has described involution as a Chinese problem and skill mismatch as a Malaysian one. The multi-group model can show whether each driver operates in both regions with different strength, or whether each is confined to one setting. A finding that involution also predicts disengagement in the Klang Valley would suggest that credential competition belongs to mass higher education systems in the region generally, not to China alone. A finding that mismatch predicts disengagement equally in both regions would weaken the contrast on which the regional framing of the research questions rests.

Fourth, the forecasting component tests whether structural variables improve projections of youth disengagement, a claim that is often made and seldom checked. Out-of-sample accuracy is the decision criterion, and differences are tested formally.

5.2 Policy Relevance

The scenario simulations are designed to inform ex-ante intervention. In Shandong, if involution is the stronger driver of long disengagement, policy attention would shift toward reducing the concentration of graduate ambitions on a small number of examination-based routes, for

example through clearer pathways into private sector and vocational employment and through recognition of professional skills alongside degrees. In the Klang Valley, if mismatch is the stronger driver, attention would shift toward the alignment of program supply with sectoral demand, starting wages in formal graduate jobs, and social protection and training access for young platform workers. In both regions, evidence that duration predicts distress and lower expectations would support early contact with young people who leave education without a job, before disengagement becomes long term. These implications are conditional on the results and will be stated as such.

5.3 Limitations

The survey uses non-probability sampling, and although quotas and weights reduce bias on observed characteristics, they cannot remove bias on unobserved ones. Stream B describes school leavers in two regions, not their whole youth populations. The survey is cross-sectional with retrospective items, so recall error is possible, especially for starting reservation wages, and causal interpretation is limited despite the sensitivity analyses. The secondary data are asymmetric: the Klang Valley series is annual and depends on access to Labour Force Survey microdata, while the Shandong series is biennial, ends in 2022 and rests on model-based small area estimates. This limits forecasting precision, most of all for Shandong, and is why panels are pooled. The Klang Valley boundary follows local authority definitions, whereas some official statistics cover Selangor as a whole, so state-level figures include districts outside the study area. Measurement invariance may hold only partially, which would restrict some comparisons. The two units are also of different kinds. Shandong is a province with large rural areas, while the Klang Valley is a metropolitan region. Urban or rural residence is controlled in every model, and the main models are re-estimated on Shandong respondents living in urban districts as a check. Finally, the two regions were chosen because they represent contrasting structural conditions, so the

results should not be generalized to other parts of China or Malaysia without further testing.

6. Conclusion

Youth employment problems in Shandong and the Klang Valley arise from different structural conditions, involution-driven competition in one case and skill mismatch with gig absorption in the other, but both produce disengagement that can leave lasting effects. This article has presented a common design for analyzing and forecasting youth employment in the two regions. Its main

practical lesson is that comparison has to be built, not read off published tables. Neither region has an official youth NEET series, so one is assembled from survey microdata in the Klang Valley and from small area estimates in Shandong, and the survey measures the drivers attributed to each region in both. The same approach can be used in other emerging economies where mass higher education and changing labor demand are reshaping the transition from school to work and where regional youth statistics are thin.

References

- Abu Rahim, M. A. R., & Sivashanmugam, L. (2024, February 23). *Underemployment and overqualification among fresh graduates* [Views]. Khazanah Research Institute. <https://www.krinstitute.org/publications/underemployment-and-overqualification-among-fresh-graduates>
- Adler, N. E., Epel, E. S., Castellazzo, G., & Ickovics, J. R. (2000). Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy, White women. *Health Psychology, 19*(6), 586–592. <https://doi.org/10.1037/0278-6133.19.6.586>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological), 57*(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Bischof, S. (2024). Test-based measurement of skill mismatch: A validation of five different measurement approaches using the NEPS. *Journal for Labour Market Research, 58*(1), Article 11. <https://doi.org/10.1186/s12651-024-00370-1>
- Boer, D., Hanke, K., & He, J. (2018). On detecting systematic measurement error in cross-cultural research: A review and critical reflection on equivalence and invariance tests. *Journal of Cross-Cultural Psychology, 49*(5), 713–734. <https://doi.org/10.1177/0022022117749042>
- Braun, V., & Clarke, V. (2022). *Thematic analysis: A practical guide*. SAGE.
- Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology, 1*(3), 185–216. <https://doi.org/10.1177/135910457000100301>
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Harvard University Press.
- Caniglia, M. (2024). Nonparticipation in work and education in emerging adulthood and depressive symptoms through early midlife. *SSM - Mental Health, 6*, 100329. <https://doi.org/10.1016/j.ssmmh.2024.100329>
- Chen, B. H. Y., Zhou, Y., & Yang, L. (2026). The role of involution on Chinese youth pursuit of postgraduate education abroad: Examine through the lens of normative social behaviour. *Asia Pacific Journal of Education, 46*(3), 905–918. <https://doi.org/10.1080/02188791.2024.2408394>
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464–504. <https://doi.org/10.1080/10705510701301834>
- Chen, X., & Chan, R. C. H. (2025). Loss of material, personal, and social resources and their impact on mental health among young people not in employment, education, or training. *Journal of Adolescence, 97*(8), 2104–2114. <https://doi.org/10.1002/jad.70024>
- Cheung, G. W., Cooper-Thomas, H. D., Lau, R. S., & Wang, L. C. (2024). Reporting reliability, convergent and discriminant validity with structural equation modeling: A review and best-practice recommendations. *Asia Pacific Journal of Management, 41*(2), 745–783. <https://doi.org/10.1007/s10490-023-09871-y>
- Deng, Y., English, A. S., & Li, Y. (2025). New elements of career construction for China's youth: Analyzing "lying flat" and work involution among emerging adults. *Emerging Adulthood, 13*(1), 131–145. <https://doi.org/10.1177/21676968241286182>

-
- Department of Statistics Malaysia. (n.d.). *Kawasanku* [Interactive dashboard]. OpenDOSM. Retrieved October 10, 2026, from <https://open.dosm.gov.my/dashboard/kawasanku>
- Department of Statistics Malaysia. (2022). *Key findings Population and Housing Census of Malaysia 2020: Administrative district*. <https://www.dosm.gov.my/portal-main/release-content/key-findings-population-and-housing-census-of-malaysia-2020-administrative-district>
- Department of Statistics Malaysia. (2025a). *Labour force statistics, Malaysia, September and third quarter 2025*. https://www.dosm.gov.my/uploads/release-content/file_20251110111813.pdf
- Department of Statistics Malaysia. (2025b). *Labour force survey report, Malaysia, 2024*. <https://dosm.gov.my/portal-main/release-content/labour-force-survey-2024>
- Department of Statistics Malaysia. (2025c). *Principal statistics of labour force, Malaysia, third quarter 2025*. <https://www.dosm.gov.my/portal-main/release-content/principal-statistics-of-labour-force-malaysia-third-quarter-2025>
- Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3), 253–263. <https://doi.org/10.1080/07350015.1995.10524599>
- Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2024). GPTs are GPTs: Labor market impact potential of LLMs. *Science*, 384(6702), 1306–1308. <https://doi.org/10.1126/science.adj0998>
- Fabrizi, E., & Rocca, A. (2024). NEET status duration and socio-economic background. *Socio-Economic Planning Sciences*, 95, 101986. <https://doi.org/10.1016/j.seps.2024.101986>
- Fay, R. E., III, & Herriot, R. A. (1979). Estimates of income for small places: An application of James-Stein procedures to census data. *Journal of the American Statistical Association*, 74(366a), 269–277. <https://doi.org/10.1080/01621459.1979.10482505>
- Felten, E., Raj, M., & Seamans, R. (2021). Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal*, 42(12), 2195–2217. <https://doi.org/10.1002/smj.3286>
- Filomena, M., Giorgetti, I., & Picchio, M. (2025). Off to a bad start: Youth nonemployment and labor market outcomes later in life. *Empirical Economics*, 69(1), 549–579. <https://doi.org/10.1007/s00181-025-02735-y>
- Gariépy, G., Danna, S. M., Hawke, L., Henderson, J., & Iyer, S. N. (2022). The mental health of young people who are not in education, employment, or training: A systematic review and meta-analysis. *Social Psychiatry and Psychiatric Epidemiology*, 57(6), 1107–1121. <https://doi.org/10.1007/s00127-021-02212-8>
- Greifer, N., & Stuart, E. A. (2021). Matching methods for confounder adjustment: An addition to the epidemiologist's toolbox. *Epidemiologic Reviews*, 43(1), 118–129. <https://doi.org/10.1093/epirev/mxab003>
- Guetterman, T. C., & Manojlovich, M. (2024). Grand rounds in methodology: Designing for integration in mixed methods research. *BMJ Quality & Safety*, 33(7), 470–478. <https://doi.org/10.1136/bmjqs-2023-016112>
- Gunnes, M., Thaulow, K., Kaspersen, S. L., Jensen, C., & Ose, S. O. (2025). Young adults not in education, employment, or training (NEET): A global scoping review. *BMC Public Health*, 25, 3394. <https://doi.org/10.1186/s12889-025-24781-y>
- Hadam, S., Würz, N., Kreutzmann, A.-K., & Schmid, T. (2024). Estimating regional unemployment with mobile network data for functional urban areas in Germany. *Statistical Methods & Applications*, 33(1), 205–233. <https://doi.org/10.1007/s10260-023-00722-0>
- Hamid, H. A., Abu Rahim, M. A. R., Suhaimi, S. A., Wan Usamah, W. A., Suhaimi, H. H., Ahmad, Z. A., & Megat Muzafar, P. M. (2024). *Shifting tides: Charting career progression of Malaysia's skilled talents*. Khazanah Research Institute. <https://www.krinstitute.org/publications/shifting-tides-charting-career-progression-of-malaysias-skilled-talents>
- Harvey, D., Leybourne, S., & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, 13(2), 281–291. [https://doi.org/10.1016/S0169-2070\(96\)00719-4](https://doi.org/10.1016/S0169-2070(96)00719-4)
- Hayes, A. F. (2022). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (3rd ed.). The Guilford Press.
- Hennink, M., & Kaiser, B. N. (2022). Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Social Science & Medicine*, 292, 114523. <https://doi.org/10.1016/j.socscimed.2021.114523>
- Hirano, K., & Imbens, G. W. (2004). The propensity score with continuous treatments. In A. Gelman & X.-L. Meng (Eds.), *Applied Bayesian modeling and causal inference from incomplete-data perspectives* (pp. 73–84). Wiley. <https://doi.org/10.1002/0470090456.ch7>
-

- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts. <https://otexts.com/fpp3/>
- International Labour Organization. (2024). *Global employment trends for youth 2024: Decent work, brighter futures*. <https://www.ilo.org/publications/major-publications/global-employment-trends-youth-2024>
- Ji, X., & Li, H. (2026). The influence mechanism of academic involution behavior among Chinese college students: A moderated mediation analysis based on the JD-R model. *Frontiers in Psychology*, 17, 1729314. <https://doi.org/10.3389/fpsyg.2026.1729314>
- Kessler, R. C., Andrews, G., Colpe, L. J., Hiripi, E., Mroczek, D. K., Normand, S.-L. T., Walters, E. E., & Zaslavsky, A. M. (2002). Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological Medicine*, 32(6), 959–976. <https://doi.org/10.1017/S0033291702006074>
- Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., Howes, M. J., Normand, S.-L. T., Manderscheid, R. W., Walters, E. E., & Zaslavsky, A. M. (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry*, 60(2), 184–189. <https://doi.org/10.1001/archpsyc.60.2.184>
- Krueger, A. B., & Mueller, A. I. (2016). A contribution to the empirics of reservation wages. *American Economic Journal: Economic Policy*, 8(1), 142–179. <https://doi.org/10.1257/pol.20140211>
- Lan, Y., Zheng, J., & Yang, Y. (2026). The meritocracy trap revisited: The effect of education on job quality among contemporary Chinese youth. *SAGE Open*, 16(1). <https://doi.org/10.1177/21582440261422455>
- Lauri, T., Toots, A., Chevalier, T., Broka, A., & Hofäcker, D. (2025). Policy mixes and youth vulnerability in Europe: A qualitative comparative analysis of the NEET rate. *Journal of European Social Policy*, 35(2), 157–172. <https://doi.org/10.1177/09589287241311159>
- Leitgöb, H., Seddig, D., Asparouhov, T., Behr, D., Davidov, E., De Roover, K., Jak, S., Meitinger, K., Menold, N., Muthén, B., Rudnev, M., Schmidt, P., & van de Schoot, R. (2023). Measurement invariance in the social sciences: Historical development, methodological challenges, state of the art, and future perspectives. *Social Science Research*, 110, 102805. <https://doi.org/10.1016/j.ssresearch.2022.102805>
- Li, C. (2025). Mapping the relationship between perceived involution and subjective well-being in Chinese university students: A network analysis across gender and grade levels. *Frontiers in Psychology*, 16, 1628064. <https://doi.org/10.3389/fpsyg.2025.1628064>
- Liu, Y., Tu, Y., Yang, H., Gao, J., Xu, Y., & Yang, Q. (2022). Have you "involution" today—Competition psychology scale for college students. *Frontiers in Psychology*, 13, 951931. <https://doi.org/10.3389/fpsyg.2022.951931>
- Low, C. W., & Mah, P. Y. (2024). The overeducation dilemma: Graduate skill mismatch in Malaysia's labour market. *International Journal of Business, Economics and Law*, 33(1), 60–66.
- Lu, H., Wang, J., & Kong, F. (2025). Bidirectional relationship between autonomy need satisfaction and "lying flat" among Chinese youth. *Australian Journal of Psychology*, 77(1), 2558847. <https://doi.org/10.1080/00049530.2025.2558847>
- Mok, K. H., & Qian, J. (2018). Massification of higher education and youth transition: Skills mismatch, informal sector jobs and implications for China. *Journal of Education and Work*, 31(4), 339–352. <https://doi.org/10.1080/13639080.2018.1479838>
- Mühlböck, M., Kalleitner, F., Steiber, N., & Kittel, B. (2022). Scarring dreams? Young people's vocational aspirations and expectations during and after unemployment. *Social Inclusion*, 10(2), 252–264. <https://doi.org/10.17645/si.v10i2.5162>
- Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(4), 599–620. https://doi.org/10.1207/S15328007SEM0904_8
- National Bureau of Statistics of China. (2024a, January 17). 关于完善分年龄组调查失业率有关情况的说明 [Explanation of the improvement of surveyed unemployment rates by age group]. https://www.stats.gov.cn/sj/zxfb/202401/t20240117_1946641.html
- National Bureau of Statistics of China. (2024b, September 20). 有关中国劳动统计年鉴中二、就业与失业的各地区城镇登记失业人员情况（2022年）未公布问题 [Reply on the non-publication of regional registered urban unemployment for 2022 in the China Labour Statistics Yearbook]. https://www.stats.gov.cn/hd/lyzx/zxgk/202410/t20241016_1957015.html

-
- National Bureau of Statistics of China. (2026). *Urban surveyed unemployment rate by age group, monthly* [Data set]. <https://data.stats.gov.cn/english/>
- Performance Management and Delivery Unit. (2010). *Economic Transformation Programme: A roadmap for Malaysia*. Prime Minister's Department.
- Podsakoff, P. M., Podsakoff, N. P., Williams, L. J., Huang, C., & Yang, J. (2024). Common method bias: It's bad, it's complex, it's widespread, and it's not easy to fix. *Annual Review of Organizational Psychology and Organizational Behavior*, 11(1), 17–61. <https://doi.org/10.1146/annurev-orgpsych-110721-040030>
- Rahmani, H., & Groot, W. (2023). Risk factors of being a youth not in education, employment or training (NEET): A scoping review. *International Journal of Educational Research*, 120, 102198. <https://doi.org/10.1016/j.ijer.2023.102198>
- Ralston, K., Everington, D., Feng, Z., & Dibben, C. (2022). Economic inactivity, not in employment, education or training (NEET) and scarring: The importance of NEET as a marker of long-term disadvantage. *Work, Employment and Society*, 36(1), 59–79. <https://doi.org/10.1177/0950017020973882>
- Ren, X., Abdullah, H., Shaffril, H. A. M., Rahman, H. A., & Zaremozhzabich, Z. (2026). A scoping review of "tang ping" (lying flat) and mental health status among Chinese youth. *PLOS ONE*, 21(3), e0342591. <https://doi.org/10.1371/journal.pone.0342591>
- Schmillen, A., & Umkehrer, M. (2017). The scars of youth: Effects of early-career unemployment on future unemployment experience. *International Labour Review*, 156(3–4), 465–494. <https://doi.org/10.1111/ilr.12079>
- Sen, A. (1999). *Development as freedom*. Oxford University Press.
- Siew, K. T., & Sze, W. Y. (2025). Labour market imperfection: A case study of graduate underemployment in Malaysia. *International Journal of Service Management and Sustainability*, 10(2), 54–73. <https://doi.org/10.24191/ijsms.v10i2.24227>
- Tleppayev, A., & Zeinolla, S. (2025). Forecasting youth unemployment through educational and demographic indicators: A panel time-series approach. *Forecasting*, 7(3), 37. <https://doi.org/10.3390/forecast7030037>
- Tomlinson, M., & Tholen, G. (2023). Scarring effects for young people in challenging economic times: A conceptual synthesis and future policy and research agenda. *Labour and Industry*, 33(3), 308–325. <https://doi.org/10.1080/10301763.2023.2251216>
- Uchiyama, Y., Furuoka, F., & Md Akhir, M. N. (2022). Gig workers, social protection and labour market inequality: Lessons from Malaysia. *Jurnal Ekonomi Malaysia*, 56(3), 165–184. <https://doi.org/10.17576/jem-2022-5603-09>
- Uchiyama, Y., Furuoka, F., & Omar, S. A. (2025). The rise and contestation of platform capitalism: Evidence from two food delivery blackouts in Malaysia. *Journal of Industrial Relations*, 67(3), 481–508. <https://doi.org/10.1177/00221856241294107>
- VanderWeele, T. J., & Ding, P. (2017). Sensitivity analysis in observational research: Introducing the E-value. *Annals of Internal Medicine*, 167(4), 268–274. <https://doi.org/10.7326/M16-2607>
- Walther, A. (2006). Regimes of youth transitions: Choice, flexibility and security in young people's experiences across different European contexts. *Young*, 14(2), 119–139. <https://doi.org/10.1177/1103308806062737>
- Wang, Y. A., & Rhemtulla, M. (2021). Power analysis for parameter estimation in structural equation modeling: A discussion and tutorial. *Advances in Methods and Practices in Psychological Science*, 4(1), 2515245920918253. <https://doi.org/10.1177/2515245920918253>
- Wright, L., Head, J., & Jivraj, S. (2021). Heterogeneity in the association between youth unemployment and mental health later in life: A quantile regression analysis of longitudinal data from English schoolchildren. *BMJ Open*, 11(7), e047997. <https://doi.org/10.1136/bmjopen-2020-047997>
- Xie, Y., & Hu, J. (2014). An introduction to the China Family Panel Studies (CFPS). *Chinese Sociological Review*, 47(1), 3–29.
- Yasin, M. Y., Zain, M. A. B. M., & Hassan, M. H. B. (2022). Urbanization and growth of Greater Kuala Lumpur: Issues and recommendations for urban growth management. *Southeast Asia: A Multidisciplinary Journal*, 22(2), 4–19. <https://fass.ubd.edu.bn/SEA/vol22-2/Urbanization%20and%20growth%20of%20Greater%20Kuala%20Lumpur.pdf>
- Zhang, B., Zhang, Q., Yao, C., & Liu, Z. (2024). The signaling paradox: Revisiting the impacts of overeducation in the Chinese labor market. *Education Sciences*, 14(8), 900. <https://doi.org/10.3390/educsci14080900>
- Zhang, L., & Li, Z. (2022). Factor structure and longitudinal measurement invariance of the K6 among a national representative elder sample of China. *BMC Public Health*, 22, 1789. <https://doi.org/10.1186/s12889-022-14193-7>
-