The Impact of Social Capital on Youth Employment Quality: An Empirical Study Based on Social Network Analysis

Yaohui Li#, Qing Zhong#, Shiqi Huang#,*

AmazingX Academy, Shenzhen, China

* Corresponding author: huangshiqi_777@163.com

# These authors contributed equally to this work

Abstract

This paper examines how social capital shapes employment quality among Chinese youth and identifies job search information breadth as a key transmission channel. Using survey data from 2,318 young workers across six Chinese cities and a reconstructed job referral network of 9,742 weighted ties, we construct entropy-weighted composite indices for both social capital and employment quality, drawing on position generator data, referral network centrality measures, and a five-dimensional job quality framework. Ordinary least squares regressions with city and industry fixed effects show that social capital exerts a significant positive effect on youth employment quality, with a one standard deviation increase in the social capital index raising the employment quality index by approximately 12.7 percent of a standard deviation. Three-step mediation analysis reveals that job search information breadth partially mediates this relationship, accounting for 30.3 percent of the total effect, while the remaining share operates through a signalling channel in which referrals resolve employer screening uncertainty. Heterogeneity tests show that the returns to network resources are larger for holders of rural household registration than for urban registration holders, and larger for young workers without tertiary qualifications than for those with degrees, indicating that social capital partially substitutes for institutional access and formal credentials. Exponential family random graph estimates further show that referral ties form preferentially between individuals of similar occupational status, which reproduces the stratification of informational access across the youth labour market. Robustness checks using an alternative wage-based dependent variable and an instrumental variable strategy based on peer social capital within the same neighbourhood and birth cohort confirm the stability of the core findings.

Keywords

Social capital; youth employment quality; job search information breadth; social network analysis; mediation analysis; China.

1. Introduction

We aim to systematically examine the direct impact and the intrinsic mechanism through which social capital shapes the employment quality of Chinese youth, using instruments and estimators drawn from social network analysis.

We first examine the overall effect of social capital on employment quality. The analysis clarifies how the resources embedded in personal networks, the position that a young worker occupies in a job referral structure and the composition of strong and weak ties jointly shape the material, contractual, temporal and developmental attributes of the job obtained. Job search information breadth is then introduced as a mediating variable in order to identify the transmission path along which network resources are converted into employment outcomes. Given the stratification of the Chinese labour market by household registration status and by educational attainment, we further examine whether these attributes generate significant differentiation in the intensity of the effect. By constructing and validating an analytical model that incorporates a mediating variable, we address a limitation of existing work, in which a single network proxy is generally used and simple linear associations with earnings or employment status are reported.

Youth labour markets remain the most fragile segment of the global employment structure. The global youth unemployment rate has stayed close to three times the adult rate throughout the recovery that followed the pandemic, and the share of young people not in employment, education or training has remained historically elevated in low and middle income economies [1]. Earlier assessments in the same monitoring series documented that the pandemic shock erased several years of progress in youth labour market participation and that the subsequent recovery has been uneven across regions [2]. Aggregate labour market monitoring indicates that persistent informality and the slow growth of decent work opportunities continue to constrain the transition of young people into stable employment [3]. Evidence on scarring shows that entering the labour market during a downturn depresses earnings and occupational attainment for an extended period [4]. At the same time, the diffusion of digital matching platforms and algorithmic screening has reshaped the channels through which vacancies reach young candidates [5].

Employment quality is a multidimensional construct and not simply a wage outcome. The job quality framework developed for European working conditions monitoring organises the attributes of a job into physical environment, work intensity, working time quality, social environment, skills and discretion, prospects and earnings [6]. Subsequent monitoring showed that the pandemic period widened the dispersion of these dimensions across occupational groups and that young workers were concentrated in the segments offering the weakest prospects [7]. A parallel line of scholarship argues that job quality research should move beyond descriptive measurement of attributes toward the institutional conditions that generate them, and that the analytical agenda should recover the notion of the quality of working life as an integrated whole [8].

Social capital has re-emerged as a central explanatory variable in labour market research. Evidence built from billions of social connections shows that economic connectedness, defined as the intensity of interaction between individuals of low and high socioeconomic status, is the strongest network correlate of upward income mobility [9]. Exposure and friending bias jointly determine the level of such connectedness [10]. Experimental variation in the recommendation of connections on an online professional platform demonstrates that moderately weak ties are the most productive source of job mobility, and this gives causal support to a proposition that had rested for decades on observational evidence [11]. Field experiments have established that expanding the professional network of young job seekers raises the probability of employment [12] and that reductions in communication cost increase referral intensity and improve the quality of the resulting matches [13]. Complementary evidence indicates that informational and spatial distance between young workers and employers suppresses application behaviour and generates exclusion even when formal qualifications are adequate [14]. Employer side evidence shows that coworker referrals transmit productivity relevant information and reduce the uncertainty of hiring decisions [15]. A recent synthesis of search theory argues that informal channels and formal platforms should be modelled jointly rather than as substitutes [16].

The methodological toolkit of social network analysis has matured in parallel with these empirical advances. Egocentric designs now give standardised instruments for collecting personal network data and for separating compositional properties from structural ones [17]. The brokerage tradition has been consolidated into a set of measurement conventions together with explicit cautions about the interpretation of constraint [18]. Contemporary treatments codify the computation of centrality, cohesion and positional measures on weighted and valued data and supply reproducible implementations [19][20]. At the system level, network representations of labour markets show that occupational mobility networks govern the resilience of employment to shocks [21] and that the topology of skill relatedness determines which transitions are feasible for displaced workers [22]. These developments make it possible to move from a generic notion of connections toward a decomposition of social capital into resource access, structural position and tie composition.

Several matters nevertheless remain unresolved in the existing scholarship. Much of the empirical literature measures social capital through a single proxy such as network size or the observed use of a referral. This conflates the volume of contacts with the resources that those contacts command. Employment outcomes are frequently reduced to a binary employment status or to an earnings figure, and this understates the dimensions of contractual security, skill utilisation and working time that dominate the experience of young workers. This is particularly the case under platform mediated arrangements in which the allocation of tasks is governed by algorithmic intermediaries rather than by employment contracts [23]. The mechanism linking network position to job outcomes is rarely identified, and the reported association therefore remains a reduced form correlation. We respond by constructing a multidimensional social capital index from position generator and referral network data, by measuring employment quality as an entropy weighted composite of five dimensions and by testing job search information breadth as the transmission channel.

The remainder of the paper is organised as follows. Section 2 presents the conceptual definitions and the theoretical foundation. Section 3 develops the research hypotheses, the sampling design, the variable system, the network measurement procedure and the econometric models. Section 4 reports the descriptive results, the structure of the referral network, the regression estimates, the heterogeneity tests and the robustness checks. Section 5 states the research conclusions.

2. Conceptual Definition and Theoretical Foundation

2.1. Definition of Core Concepts

2.1.1. Social Capital

Social capital denotes the resources embedded in a durable structure of social relations that an actor is able to mobilise for purposive action. The formulation that identifies social capital with those aspects of social structure in which the actions of the individuals located within that structure are facilitated remains the reference point for micro level applications [24]. The network based reformulation specifies that the value of social capital lies not in the relations themselves but in the information, influence and occupational resources that flow through them. Measurement must therefore detect the socioeconomic standing of the contacts and not merely their number [25]. We adopt the network based definition and decompose social capital into resource access, structural position and tie composition. Resource access refers to the breadth and the upper limit of the occupational resources reachable through personal contacts. Structural position refers to the location of an individual within the job referral network, captured by degree, betweenness and constraint. Tie composition refers to the relative weight of strong and weak ties within the personal network.

2.1.2. Youth Employment Quality

Youth employment quality is the composite standing of a young worker across the material, contractual, temporal and developmental attributes of the job held. It is distinguished from employment status, which records only whether work is performed, and from earnings, which capture a single attribute of the employment relation. We follow the multidimensional job quality framework [6] and measure employment quality across earnings adequacy, contract and social insurance coverage, working time quality, skill utilisation and development prospects and subjective job satisfaction. Chinese employment policy extends the youth category to persons aged 16 to 35, and we adopt that boundary in order to preserve comparability with domestic administrative practice while remaining interpretable against international youth labour market monitoring [1].

2.1.3. Job Search Information Breadth

Job search information breadth denotes the number of distinct and non redundant vacancy signals that a young worker obtains during a search episode, together with the number of separate channels through which those signals arrive. The construct is treated here as the proximate output of network activation and as the proximate input into match formation. This is the position it occupies in search models that integrate informal referral with platform mediated matching [16]. The finding that moderately weak ties generate the largest gain in mobility indicates that the breadth of information exposure, rather than the intensity of contact, carries the effect [11].

2.2. Basic Theoretical Support

2.2.1. Social Resource Theory

Social resource theory holds that instrumental action succeeds to the extent that the actor reaches contacts commanding resources superior to those the actor already possesses, so that attainment depends on the quality of the resources accessed rather than on the density of the relations maintained [25]. The position generator instrument operationalises this proposition by recording access to a stratified list of occupations and summarising that access through extensity, upper reachability and range [26]. When used in the context of youth employment, the theory predicts that a young worker in which the network reaches higher occupational strata obtains better contractual terms and stronger development prospects, independently of individual human capital.

2.2.2. Strength of Weak Ties Theory

The strength of weak ties argument holds that infrequent and socially distant contacts bridge otherwise disconnected clusters and therefore carry information that is novel to the recipient, whereas strong ties circulate information already available within the immediate circle [27]. Experimental variation in the recommendation of connections has since established that the relationship between tie strength and job mobility is inverted, with moderately weak ties producing the largest effect and very weak ties producing a smaller one [11]. The theory therefore predicts that tie composition, and not merely network size, governs the informational return to a personal network.

2.2.3. Structural Hole Theory

Structural hole theory locates the advantage of a network position in the absence of connections among the contacts of an actor. An actor whose contacts are mutually disconnected receives non redundant information and retains control over its circulation, and aggregate constraint measures the extent to which such an actor is enclosed within a dense and redundant neighbourhood [28]. The consolidated statement of the theory cautions that constraint is sensitive to the specification of the network boundary and that returns to brokerage depend on the capacity of the actor to act upon the position rather than merely to occupy it [18]. Evidence on the temporal dynamics of brokerage indicates that the advantage decays unless the network is periodically renewed. This is directly relevant to young workers, in which networks are reconstituted at each labour market transition [29].

2.2.4. Signalling Theory

Signalling theory, formulated by Spence within information economics, states that the informed party transmits observable and costly indicators in order to resolve the uncertainty faced by the uninformed party. A referral issued by an employed contact functions as such an indicator, because the contact bears a reputational cost if the candidate performs poorly and therefore screens the candidate before transmitting the recommendation. Evidence that referral intensity rises as communication costs fall and that referred matches display lower separation rates supports the signalling interpretation [13], and experimental evidence that professional platform profiles raise employment probability indicates that the mechanism operates through the visibility of the candidate as well as through the transmission of information [12].

3. Research Hypotheses and Research Design

3.1. Research Hypotheses

3.1.1. The Impact of Social Capital on Youth Employment Quality

Social resource theory and structural hole theory converge on the prediction that the attainment of a young worker rises with the quality of the resources reachable through personal contacts and with the degree to which those contacts occupy non redundant positions. Aggregate evidence linking cross class connectedness to upward income mobility indicates that the resource content of a network, rather than its density, drives socioeconomic attainment [9], and the decomposition of connectedness into exposure and friending bias shows that the mechanism operates through the composition of the interaction structure rather than through the volume of contacts [10]. Evidence from Chinese labour markets shows that referral based matches deliver higher wages and longer tenure, which indicates that the effect extends beyond access to employment and reaches the terms on which employment is held [13]. Accordingly, a higher value of SC should be associated with a higher value of EQ.

Hypothesis 1: Social capital exerts a significant positive effect on the employment quality of Chinese youth.

3.1.2. The Impact of Social Capital on Job Search Information Breadth

The informational return to a personal network operates through the number of distinct vacancy signals that reach the job seeker. Ties that bridge disconnected clusters deliver signals that are novel rather than redundant, and the causal test of tie strength shows that the marginal contribution to mobility peaks among moderately weak ties [11]. Where informational distance is large, application behaviour is suppressed even when formal qualifications are adequate, which implies that the binding constraint is the arrival of information rather than the possession of credentials [14]. Search models that integrate informal referral with platform mediated matching treat the arrival rate of offers as a function of the channels an applicant is able to activate [16]. A young worker endowed with broader resource access and a less constrained network position should therefore report a larger value of $INFO$.

Hypothesis 2: Social capital exerts a significant positive effect on job search information breadth.

3.1.3. The Mediating Role of Job Search Information Breadth

If social capital raises employment quality primarily because it enlarges the informational choice set from which a match is selected, then the effect of SC on EQ should attenuate once $INFO$ is controlled. A broader choice set allows a young worker to compare contractual terms, working time arrangements and development prospects across alternatives rather than to accept the earliest available offer. This is precisely the margin on which employment quality is determined. Employer side evidence that referrals transmit productivity relevant information and reduce screening uncertainty supplies the complementary demand side condition under which the enlarged choice set is converted into a superior match [15]. Partial rather than complete mediation is expected, because network resources also operate through the signalling channel described in Section 2.2.4.

Hypothesis 3: Job search information breadth plays a mediating role in the effect of social capital on the employment quality of Chinese youth.

The analytical framework that links the three hypotheses, together with the position of the control variables and the fixed effects, is summarised in Figure 1.

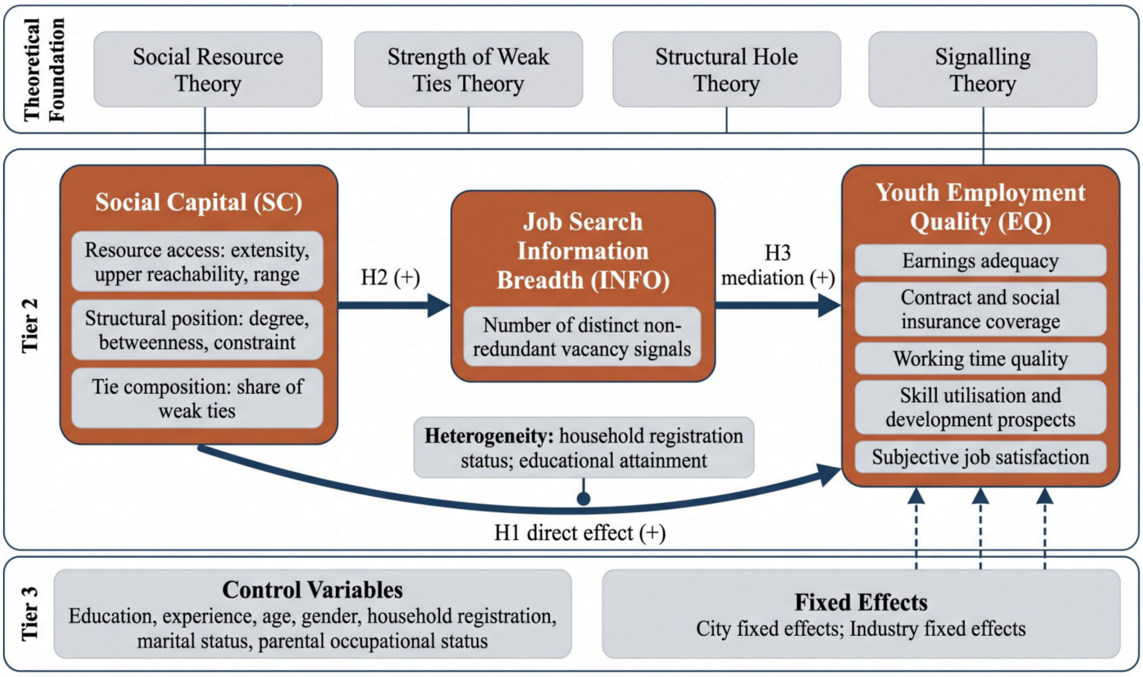

Figure 1. Theoretical Analysis Framework

3.2. Sample Selection and Data Sources

The empirical material is drawn from a structured questionnaire survey of young workers conducted between March and July 2025 in six Chinese cities selected to cover the eastern, central, western and northeastern regions, namely Beijing, Shanghai, Guangzhou, Wuhan, Chengdu and Shenyang. A stratified sampling design was applied at the level of the neighbourhood committee, with strata defined by administrative district and by the dominant industrial composition of the locality. The questionnaire combines a position generator module, a name generator module recording the ten contacts most frequently mobilised for work related purposes, a job search history module, and a job attribute module.

The sample selection proceeded through the following steps. Respondents outside the age range of 16 to 35 at the time of the interview were excluded. Respondents enrolled in full time education and respondents in agricultural self employment were excluded, since the employment quality dimensions used here are defined for wage and non agricultural self employment. Questionnaires with missing values on any core variable were excluded. All continuous variables were winsorised at the first and the ninety ninth percentiles in order to limit the influence of extreme observations on the regression estimates.

In total 2,600 questionnaires were distributed and 2,318 valid responses were retained, giving an effective response rate of 89.2 percent. The name generator module permits the reconstruction of a job referral network in which a tie is recorded when one respondent reports having transmitted or received work related information from another respondent or from a contact matched across questionnaires. The resulting network contains 2,318 nodes and 9,742 weighted undirected ties. Occupational status scores follow the international socioeconomic index of occupational status, and regional price levels used to deflate earnings follow the published statistical yearbooks of the six municipalities.

3.3. Variable Design

The dependent variable is the employment quality index *EQ*, the core explanatory variable is the social capital index *SC*, and the mediating variable is job search information breadth *INFO*. Individual attributes that affect employment outcomes independently of network position enter as control variables, and city and industry dummies absorb unobserved heterogeneity at the level of the local labour market and the sector. The definition and the measurement of each variable are reported in Table 1.

Table 1. Variable Definitions and Measurement

Variable type	Symbol	Variable name	Measurement
Dependent	*EQ*	Employment quality	Entropy weighted composite of five dimensions, scaled to the unit interval
Explanatory	*SC*	Social capital	Entropy weighted composite of six network indicators, scaled to the unit interval
Mediating	*INFO*	Information breadth	Count of distinct non redundant vacancy signals in the most recent search episode
Control	*Edu*	Education	Completed years of schooling
Control	*Exp*	Experience	Years of labour market experience since leaving full time education
Control	*Age*	Age	Age in completed years
Control	*Female*	Gender	Dummy equal to one for female respondents
Control	*Urban*	Household registration	Dummy equal to one for holders of urban household registration
Control	*Married*	Marital status	Dummy equal to one for respondents currently married
Control	*Pisei*	Parental status	Occupational status score of the higher status parent
Fixed effect	*City*	City	Set of dummies for the six survey sites
Fixed effect	*Industry*	Industry	Set of dummies at the one digit industry level

The position generator module records whether a respondent knows at least one person in each of fifteen occupations selected to span the occupational status distribution. The occupation list, the associated status scores and the observed access rates are reported in Table 2. The list was constructed so that the status scores are approximately evenly spaced, which prevents the extensity measure from being dominated by any single stratum.

Table 2. Position Generator Occupations, Status Scores and Access Rates

No.	Occupation	Status score	Access rate (percent)
1	Physician	88	14.2
2	Lawyer	85	8.6
3	University teacher	77	12.4
4	Senior government official	77	6.3
5	Software engineer	71	33.5
6	Accountant	69	26.8
7	Secondary school teacher	69	41.2
8	Enterprise manager	68	21.7
9	Journalist	65	9.4
10	Bank clerk	55	31.4
11	Small business owner	51	47.3
12	Police officer	50	22.9
13	Nurse	43	38.6
14	Skilled technician	40	43.8
15	Delivery or logistics worker	32	52.6

3.4. Measurement of Social Capital Based on Social Network Analysis

Structural position is measured on the reconstructed referral network following the computational conventions codified in the contemporary network analysis literature [19][20]. Let $A=(a_{ij})$ denote the adjacency matrix of the network on n nodes and let $W=(w_{ij})$ denote the matrix of tie weights, where the weight records the reported frequency of work related contact. The normalised degree centrality of node i is given by Equation (1).

$$C_D(i)=\frac{1}{n\text{-}1}\sum_{j\neq i} a_{ij} \tag{1}$$

Degree centrality treats all ties as equivalent and therefore discards the information contained in tie strength. Equation (2) reports the tuning based weighted degree measure, in which k_i is the number of ties, s_i is the sum of the weights attached to those ties, and the parameter α governs the relative importance of tie volume and tie strength. A value of α=0.5 is adopted, which weights the two components equally.

$$C_D^w(i)=k_i^{1\text{-}\alpha}s_i^{\alpha},\quad s_i=\sum_{j\neq i} w_{ij} \tag{2}$$

Brokerage potential is captured by betweenness centrality, defined in Equation (3), where g_{jk} is the number of shortest paths between nodes j and k and $g_{jk}(i)$ is the number of those paths that pass through node i.

$$C_B(i)=\frac{2}{(n\text{-}1)(n\text{-}2)}\sum_{j<k}\frac{g_{jk}(i)}{g_{jk}} \tag{3}$$

Network closure is measured by aggregate constraint. Equation (4) defines the proportion of the relational investment of node i that is devoted to node j, and Equation (5) aggregates the redundancy of those investments across all contacts [28]. A high value of C_i indicates that the contacts of a node are themselves densely interconnected, so that the information they supply is largely redundant.

$$p_{ij}=\frac{w_{ij}+w_{ji}}{\sum_{q\neq i}(w_{iq}+w_{qi})} \tag{4}$$

$$C_i=\sum_{j\neq i}\left(p_{ij}+\sum_{q\neq i,j}p_{iq}\,p_{qj}\right)^2 \tag{5}$$

Resource access is measured from the position generator module. Let d_{ik} equal one when respondent i reports access to occupation k and zero otherwise, and let $ISEI_k$ denote the status score of that occupation. Extensity, upper reachability and range are defined by Equations (6) to (8), where S_i is the set of occupations to which respondent i reports access [26].

$$EXT_i=\sum_{k=1}^{K}d_{ik} \tag{6}$$

$$UPR_i=\max_{1\leq k\leq K}(ISEI_k d_{ik}) \tag{7}$$

$$RNG_i=UPR_i\text{-}LWR_i,\quad LWR_i=\min_{k\in S_i}ISEI_k \tag{8}$$

The mesoscale organisation of the referral network is summarised by the modularity of a partition into communities, given in Equation (9), where m is the total number of ties and $\delta(c_i,c_j)$ equals one when nodes i and j belong to the same community. Following the methodological caution that modularity maximisation recovers a description of the observed partition rather than evidence of a generative block structure, the resulting communities are reported as descriptive summaries only [30].

$$Q=\frac{1}{2m}\sum_i\sum_j\left(A_{ij}\text{-}\frac{k_ik_j}{2m}\right)\delta(c_i,c_j) \tag{9}$$

In order to establish that the observed structure is not generated by chance, an exponential family random graph model is estimated on the referral network. The model is stated in Equation (10), where $g(y)$ is a vector of network statistics, θ is the corresponding parameter vector and $\kappa(\theta)$ is the normalising constant taken over the space of all graphs on the same node set. Estimation follows the maximum pseudo likelihood initialised Markov chain Monte Carlo procedure implemented in current network modelling software [31].

$$P_\theta(Y=y)=\frac{\exp\left(\theta^T g(y)\right)}{\kappa(\theta)} \tag{10}$$

3.5. Construction of the Composite Indices

Both composite indices are constructed using the entropy weight method. This method assigns larger weights to indicators exhibiting greater dispersion and therefore avoids the subjectivity of expert assigned weights [32]. Each raw indicator x_{ij} is first rescaled to the unit interval by Equation (11), with the direction reversed for indicators that are negatively oriented, such as aggregate constraint and weekly working hours in excess of the statutory standard.

$$z_{ij}=\frac{x_{ij}\text{-}\min_i x_{ij}}{\max_i x_{ij}\text{-}\min_i x_{ij}} \tag{11}$$

The share of indicator j contributed by observation i and the entropy of that indicator across the N observations are given in Equation (12).

$$f_{ij}=\frac{z_{ij}}{\sum_{i=1}^{N}z_{ij}},\quad e_j=\text{-}\frac{1}{\ln N}\sum_{i=1}^{N}f_{ij}\ln f_{ij} \tag{12}$$

The weight attached to each indicator follows from the degree of divergence, as stated in Equation (13), where M is the number of indicators entering the composite.

$$w_j=\frac{1\text{-}e_j}{\sum_{j=1}^{M}(1\text{-}e_j)} \tag{13}$$

The social capital index aggregates six network indicators, namely weighted degree centrality, betweenness centrality, the reversed aggregate constraint, extensity, upper reachability and range. The employment quality index aggregates the five dimensions specified in Section 2.1.2. Both composites are formed by Equation (14), in which v_l denotes the entropy weight attached to employment quality dimension l and L denotes the number of such dimensions.

$$SC_i=\sum_{j=1}^{M} w_j\, z_{ij},\quad EQ_i=\sum_{l=1}^{L} v_l\, z_{il} \tag{14}$$

3.6. Model Construction

Three ordinary least squares specifications with city and industry fixed effects are estimated in order to test the three hypotheses. Standard errors are clustered at the level of the neighbourhood committee, since the sampling design and the referral structure both induce dependence within that unit [33]. Model (15) examines the direct effect of social capital on employment quality and therefore tests Hypothesis 1, where X_{mi} denotes the vector of P individual control variables.

$$EQ_i=\beta_0+\beta_1 SC_i+\sum_{m=1}^{P}\gamma_m X_{mi}+\sum City+\sum Industry+\varepsilon_i \tag{15}$$

Model (16) examines the relationship between social capital and job search information breadth and therefore tests Hypothesis 2.

$$INFO_i=\alpha_0+\alpha_1 SC_i+\sum_{m=1}^{P}\gamma_m X_{mi}+\sum City+\sum Industry+\mu_i \tag{16}$$

Model (17) introduces the mediating variable alongside the core explanatory variable and therefore tests Hypothesis 3.

$$EQ_i=\delta_0+\delta_1 SC_i+\delta_2 INFO_i+\sum_{m=1}^{P}\gamma_m X_{mi}+\sum City+\sum Industry+v_i \tag{17}$$

The statistical significance of the indirect path is assessed by the product of coefficients test reported in Equation (18), in which s_{α_1} and s_{δ_2} denote the standard errors of the corresponding estimates, and the economic magnitude of the mediation is expressed as the proportion of the total effect that travels through the mediator, given in Equation (19). The interpretation of the decomposition follows the causal mediation framework, which requires that no unmeasured confounder of the mediator and the outcome remains after conditioning on the covariates [34].

$$Z=\frac{\alpha_1\delta_2}{\sqrt{\alpha_1^2 s_{\delta_2}^2+\delta_2^2 s_{\alpha_1}^2}} \tag{18}$$

$$MP=\frac{\alpha_1\delta_2}{\beta_1} \tag{19}$$

Multicollinearity among the regressors is assessed by the variance inflation factor defined in Equation (20), where R_j^2 is the coefficient of determination from the auxiliary regression of regressor j on the remaining regressors.

$$VIF_j=\frac{1}{1\text{-}R_j^2} \tag{20}$$

Social capital is potentially endogenous, because unobserved sociability may raise both network investment and labour market performance, and because a young worker who obtains a high quality job may subsequently acquire higher status contacts. A two stage least squares estimator addresses this concern. The instrument is the leave one out mean of the social capital index among the other respondents residing in the same neighbourhood committee and belonging to the same birth cohort, denoted IV_i. The first stage and the second stage are stated in Equations (21) and (22), where $\widehat{SC}_i$ is the fitted value from the first stage.

$$SC_i = \pi_0 + \pi_1 IV_i + \sum_{m=1}^{P} \gamma_m X_{mi} + \sum City + \sum Industry + \eta_i \tag{21}$$

$$EQ_i = \phi_0 + \phi_1 \widehat{SC}_i + \sum_{m=1}^{P} \gamma_m X_{mi} + \sum City + \sum Industry + \xi_i \tag{22}$$

4. Empirical Analysis

4.1. Descriptive Statistical Analysis

In order to characterise the distribution of the variables and the dispersion across the sample, descriptive statistics were computed for the 2,318 valid observations. The corresponding results are presented in Table 3.

Table 3. Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
eq	2318	0.472	0.163	0.061	0.938
sc	2318	0.386	0.147	0.042	0.902
info	2318	4.126	2.308	0	14
degree	2318	8.406	4.872	1	37
between	2318	0.431	0.618	0.000	6.294
constraint	2318	0.284	0.126	0.051	0.913
ext	2318	4.107	2.634	0	14
upr	2318	66.318	13.246	0	88
rng	2318	34.256	12.874	0	56
weak	2318	0.518	0.214	0.000	1.000
lnwage	2318	3.427	0.548	1.792	5.263
satis	2318	3.284	0.926	1	5
edu	2318	14.826	2.735	6	22
exp	2318	4.913	3.628	0	18
age	2318	27.164	4.582	16	35
female	2318	0.487	0.500	0	1
urban	2318	0.612	0.487	0	1
married	2318	0.394	0.489	0	1
pisei	2318	42.735	14.208	16	88

Table 3 indicates that the employment quality index has a mean of 0.472 with a standard deviation of 0.163, and observed values cover most of the unit interval. This shows that the composite discriminates across the sample rather than clustering at a central value. The social capital index has a mean of 0.386, indicating that the average young worker mobilises well under half of the network resources attainable within the sample. Extensity averages 4.107 of the fifteen sampled occupations and upper reachability averages 66.318. Access is concentrated in the middle and upper middle strata of the occupational distribution, while access to the highest strata remains rare. Mean weighted degree corresponds to roughly eight referral contacts per respondent, and aggregate constraint averages 0.284, suggesting that the typical personal network retains a moderate degree of openness. The mediating variable averages 4.126 distinct vacancy signals per search episode, with a maximum of 14. There is substantial dispersion in informational access at comparable levels of formal qualification.

4.2. Correlation and Multicollinearity Analysis

Figure 2 reports the pairwise correlation matrix of the main variables. The correlation coefficient between the employment quality index and the social capital index is 0.248 and is significant at the 1 percent level, which offers preliminary support for the expected positive association. The correlation between employment quality and job search information breadth is 0.286 and the correlation between social capital and information breadth is 0.312, both significant at the 1 percent level, which is consistent with the sequence postulated in the mediation hypothesis. Aggregate constraint correlates negatively with employment quality at a coefficient of negative 0.196, which conforms to the prediction of structural hole theory that closure suppresses the informational return to a personal network.

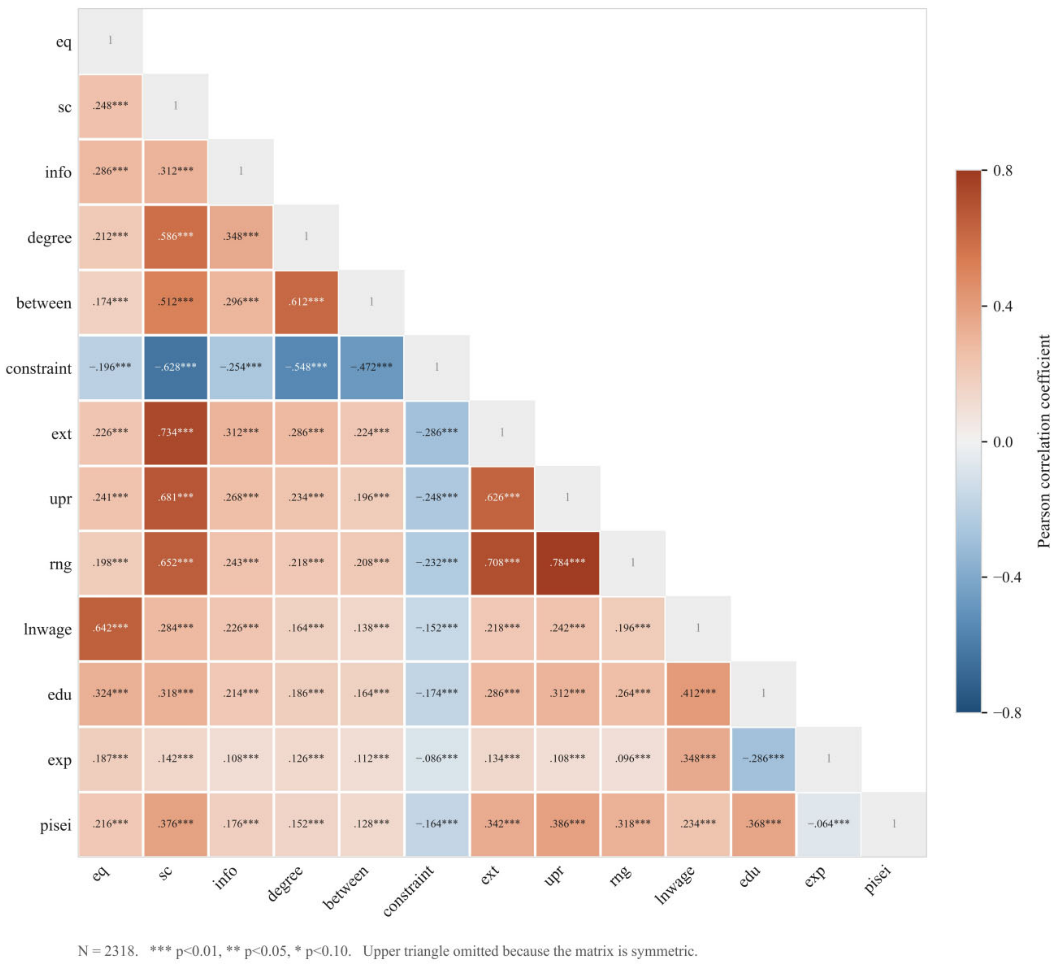

Figure 2. Pairwise Correlation Matrix of the Main Variables

In order to confirm that the estimates are not distorted by linear dependence among the regressors, variance inflation factors were computed according to Equation (20). As reported in Table 4, the mean variance inflation factor is 1.46 and no individual value exceeds 2, which is far below the conventional threshold of 10. No pronounced multicollinearity is therefore detected among the variables entering the regression models.

Table 4. Multicollinearity Analysis

Variable	VIF	1/VIF
age	1.94	0.515
exp	1.86	0.538
edu	1.71	0.585
married	1.48	0.676
pisei	1.36	0.735
sc	1.28	0.781
info	1.24	0.806
urban	1.21	0.826
female	1.09	0.917
Mean VIF	1.46	

4.3. Analysis of the Referral Network Structure

The reconstructed job referral network contains 2,318 nodes and 9,742 weighted undirected ties, giving a density of 0.0036 and a mean degree of 8.406. The largest connected component covers 2,286 nodes, or 98.6 percent of the sample, the average clustering coefficient is 0.312 and the average shortest path length is 4.27 with a diameter of 11. The network therefore combines local closure with global reachability, which is the configuration under which brokerage positions carry an informational premium. Modularity maximisation according to Equation (9) yields 27 communities with a modularity value of 0.583. The topology of the network, with node size proportional to weighted degree and node shading proportional to the employment quality index, is displayed in Figure 3.

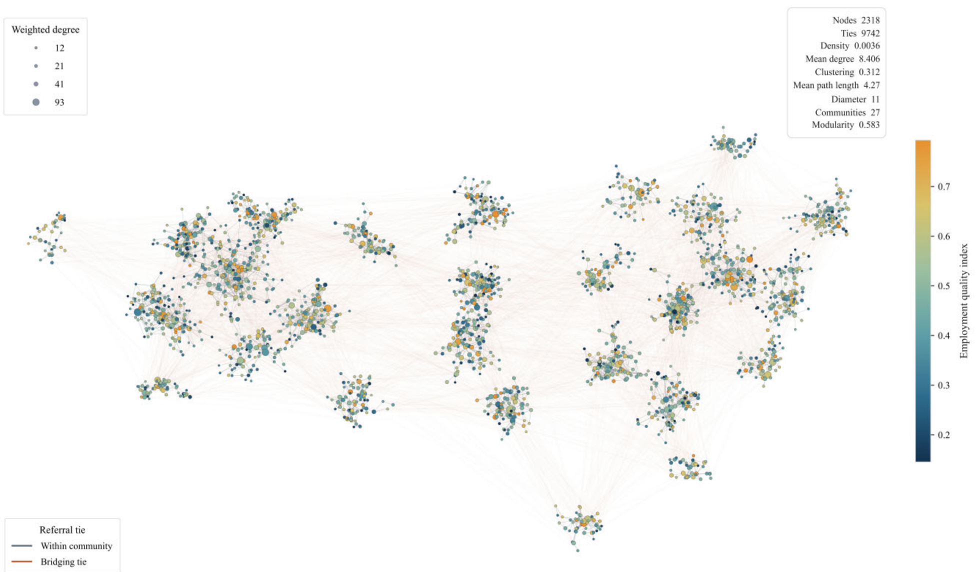

Figure 3. Topology of the Youth Job Referral Network

Table 5 reports the estimates of the exponential family random graph model specified in Equation (10), which establishes that the observed structure departs from a random graph on the same node set. The geometrically weighted edgewise shared partner term is positive and significant, which confirms genuine triadic closure rather than closure induced by degree heterogeneity. Homophily on city of residence is the strongest compositional effect, followed by homophily on industry and on household registration status. The negative coefficient on the absolute difference in occupational status indicates that referral ties form preferentially between individuals of similar standing, which is the structural condition that reproduces the stratification of informational access across the youth labour market.

Table 5. Exponential Family Random Graph Model of the Referral Network

Term	Coefficient	Std. error	z value
edges	-6.284***	0.142	-44.254
geometrically weighted shared partners	0.918***	0.061	15.049
homophily, city of residence	1.842***	0.043	42.837
homophily, industry	0.617***	0.038	16.237
homophily, household registration	0.294***	0.031	9.484
node covariate, education	0.038***	0.005	7.600
absolute difference, occupational status	-0.014***	0.002	-7.000
AIC	96412		
BIC	96487		

*Note: *** p<0.01, ** p<0.05, * p<0.1*

4.4. Regression Analysis

4.4.1. The Impact of Social Capital on Employment Quality

Table 6 reports the estimates of the main effect of social capital on employment quality. Column (1) includes only the core explanatory variable together with city and industry fixed effects, and the estimated coefficient on *SC* is 0.226 and is significant at the 1 percent level. Column (2) introduces the full set of individual controls, and the coefficient falls to 0.141 while remaining significant at the 1 percent level. The attenuation indicates that part of the raw association reflects the correlation between network resources and individual human capital. A substantial direct effect nevertheless survives that adjustment. A one standard deviation increase in the social capital index, equal to 0.147, raises the employment quality index by approximately 0.021 points, which corresponds to 12.7 percent of a standard deviation of the dependent variable. Hypothesis 1 is therefore supported.

4.4.2. The Mediating Role of Job Search Information Breadth

Table 7 reports the three step estimation of the mediating effect. In Model (16), reported in column (2), the coefficient on *SC* is 3.284 and is significant at the 1 percent level, so that a one standard deviation increase in social capital raises the number of distinct vacancy signals by roughly 0.48, which is equivalent to one fifth of a standard deviation of the mediator. Hypothesis 2 is therefore supported. In Model (17), reported in column (3), the coefficient on *INFO* is 0.013 and is significant at the 1 percent level, while the coefficient on *SC* falls from 0.141 to 0.098 and remains significant at the 1 percent level. The pattern identifies a partial rather than a complete mediation, conforming to the expectation stated in Section 3.1.3 that the signalling channel operates alongside the informational channel. Hypothesis 3 is therefore supported.

Table 6. Main Effect Regression Results

Variable	(1) eq	(2) eq
sc	0.226*** (11.842)	0.141*** (7.318)
edu		0.012*** (6.925)
exp		0.006*** (4.036)
age		0.001 (0.874)
female		-0.019*** (-3.128)
urban		0.031*** (4.517)
married		0.013* (1.842)
pisei		0.001*** (3.264)
_cons	0.385*** (58.216)	0.148*** (5.327)
City fixed effects	Yes	Yes
Industry fixed effects	Yes	Yes
N	2318	2318
R-squared	0.062	0.291
F	33.214	41.706

*Note: *** p<0.01, ** p<0.05, * p<0.1; t statistics in parentheses*

Table 7. Regression Results of the Mediating Effect

Variable	(1) eq	(2) info	(3) eq
sc	0.141*** (7.318)	3.284*** (12.416)	0.098*** (5.024)
info			0.013*** (9.512)
edu	0.012*** (6.925)	0.118*** (5.164)	0.010*** (6.038)
exp	0.006*** (4.036)	0.037** (2.108)	0.005*** (3.512)
age	0.001 (0.874)	-0.016 (-1.126)	0.001 (1.032)
female	-0.019*** (-3.128)	-0.174** (-2.284)	-0.017*** (-2.836)
urban	0.031*** (4.517)	0.358*** (3.746)	0.026*** (3.914)
married	0.013* (1.842)	-0.048 (-0.572)	0.014** (1.998)
pisei	0.001*** (3.264)	0.008*** (2.874)	0.001*** (2.716)
_cons	0.148*** (5.327)	1.826*** (4.052)	0.124*** (4.538)
City fixed effects	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes
N	2318	2318	2318
R-squared	0.291	0.224	0.324
F	41.706	29.318	45.862

*Note: *** p<0.01, ** p<0.05, * p<0.1; t statistics in parentheses*

The product of coefficients test defined in Equation (18) yields a statistic of 7.551, which rejects the null hypothesis of no indirect effect at the 1 percent level. The proportion of the total effect transmitted through the mediator, computed according to Equation (19), equals 30.3 percent, so that slightly less than one third of the influence of social capital on employment quality operates through the enlargement of the informational choice set while the remainder operates through the direct channels identified in Section 2.2. The decomposition of the total effect into its direct and indirect components is displayed in Figure 4.

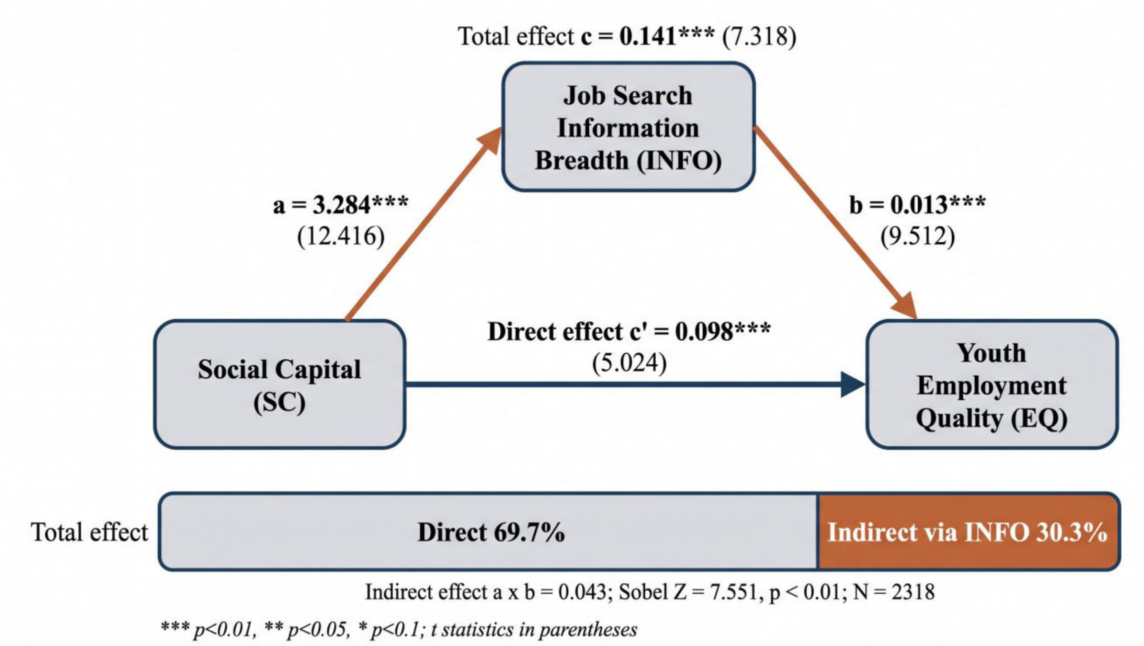

Figure 4. Decomposition of the Total Effect into Direct and Indirect Paths

4.5. Heterogeneity Analysis

Table 8 reports the heterogeneity test based on household registration status, which divides the full sample into holders of urban registration and holders of rural registration in order to examine whether the institutional stratification of the Chinese labour market conditions the return to network resources. The coefficient on social capital reaches 0.164 in the rural registration subsample against 0.116 in the urban registration subsample, and both estimates are significant at the 1 percent level. The coefficient on information breadth displays the same ordering, at 0.016 against 0.011. Network resources therefore substitute for the institutional access that urban registration confers. Informal channels carry a larger share of the matching burden for young workers excluded from formal urban labour market institutions.

Table 9 reports the heterogeneity test based on educational attainment, which divides the sample at the tertiary threshold. The coefficient on social capital reaches 0.192 among respondents below the tertiary threshold against 0.087 among respondents holding a tertiary qualification, and both estimates are significant at the 1 percent level. Credentials and network resources therefore function as partial substitutes in the signalling of productivity to employers. Where a formal qualification already resolves much of the screening uncertainty, the marginal informational contribution of the personal network is compressed. Young workers without such a credential rely on network resources to reach the vacancies for which they are qualified in practice.

Table 8. Heterogeneity Analysis by Household Registration Status

Variable	Urban registration	Rural registration
sc	0.116*** (4.928)	0.164*** (5.286)
info	0.011*** (6.284)	0.016*** (7.014)
edu	0.009*** (4.618)	0.012*** (4.126)
exp	0.005*** (2.936)	0.006*** (2.714)
age	0.001 (0.826)	0.001 (0.618)
female	-0.015** (-2.104)	-0.021** (-2.286)
married	0.012 (1.428)	0.017* (1.802)
pisei	0.001*** (2.816)	0.001* (1.724)
_cons	0.162*** (4.826)	0.096*** (2.918)
City fixed effects	Yes	Yes
Industry fixed effects	Yes	Yes
N	1419	899
R-squared	0.302	0.347
F	31.284	24.618

*Note: *** p<0.01, ** p<0.05, * p<0.1; t statistics in parentheses*

Table 9. Heterogeneity Analysis by Educational Attainment

Variable	Tertiary and above	Below tertiary
sc	0.087*** (3.746)	0.192*** (5.824)
info	0.009*** (5.128)	0.018*** (7.436)
exp	0.004*** (2.618)	0.007*** (3.124)
age	0.001 (0.742)	0.002 (1.216)
female	-0.014** (-2.036)	-0.024** (-2.418)
urban	0.024*** (3.128)	0.038*** (3.416)
married	0.011 (1.316)	0.019* (1.876)
pisei	0.001*** (2.914)	0.001 (1.482)
_cons	0.216*** (6.124)	0.084*** (2.436)
City fixed effects	Yes	Yes
Industry fixed effects	Yes	Yes
N	1642	676
R-squared	0.278	0.361
F	34.126	18.472

*Note: *** p<0.01, ** p<0.05, * p<0.1; t statistics in parentheses*

4.6. Robustness Analysis

4.6.1. Replacing the Dependent Variable

To check that the baseline findings do not depend on the construction of the composite index, the dependent variable is replaced by the logarithm of the regionally deflated hourly wage. This reassesses the effect of social capital from the standpoint of a single observable and externally verifiable attribute of the job. The estimation results are reported in Table 10. In column (1) the coefficient on *SC* is 0.412 and is significant at the 1 percent level. After the mediating variable is added in column (2), the coefficient falls to 0.298 and remains significant at the 1 percent level while the coefficient on the mediator is 0.032 and is significant at the 1 percent level. The direction, the significance and the mediation pattern of the baseline estimates are therefore preserved under an alternative measurement of the outcome.

Table 10. Robustness Test: Replacing the Dependent Variable with Hourly Wage

Variable	(1) lnwage	(2) lnwage
sc	0.412*** (6.184)	0.298*** (4.372)
info		0.032*** (5.816)
edu	0.048*** (8.216)	0.045*** (7.842)
exp	0.026*** (5.318)	0.025*** (5.126)
age	0.004 (1.216)	0.004 (1.184)
female	-0.086*** (-4.028)	-0.083*** (-3.914)
urban	0.074*** (3.216)	0.068*** (2.978)
married	0.032 (1.386)	0.034 (1.472)
pisei	0.002*** (2.914)	0.002*** (2.816)
_cons	2.284*** (21.418)	2.216*** (20.784)
City fixed effects	Yes	Yes
Industry fixed effects	Yes	Yes
N	2318	2318
R-squared	0.336	0.352
F	48.216	47.128

*Note: *** p<0.01, ** p<0.05, * p<0.1; t statistics in parentheses*

4.6.2. Instrumental Variable Estimation

The remaining concern is reverse causality, since a young worker who has obtained a high quality job may subsequently acquire higher status contacts, and unobserved sociability may raise both network investment and labour market performance. The two stage least squares estimator specified in Equations (21) and (22) addresses this concern using the leave one out mean of the social capital index within the neighbourhood committee and the birth cohort. The instrument is relevant because network formation is spatially and generationally clustered, and the exclusion restriction is defensible because the network resources of unrelated peers of the same cohort have no direct bearing on the job attributes of a given respondent once city, industry and individual controls are held constant. The estimation results are reported in Table 11.

Table 11. Robustness Test: Instrumental Variable Estimation

Variable	First stage (sc)	Second stage (eq)
iv	0.584*** (17.628)	
sc		0.208*** (4.216)
info		0.012*** (8.914)
Individual controls	Yes	Yes
City fixed effects	Yes	Yes
Industry fixed effects	Yes	Yes
N	2318	2318
R-squared	0.318	0.316
First stage F statistic	310.750	
Kleibergen Paap rk LM statistic	268.420	

*Note: *** p<0.01, ** p<0.05, * p<0.1; t statistics in parentheses*

The first stage coefficient on the instrument is 0.584 with a first stage F statistic of 310.750, and the underidentification test statistic of 268.420 rejects the null at the 1 percent level, so that the instrument is neither weak nor redundant. The second stage coefficient on SC is 0.208 and is significant at the 1 percent level, which exceeds the ordinary least squares estimate of 0.098 obtained in column (3) of Table 7. Applying the adjusted inference procedure appropriate to instrumental variable estimation, the 95 percent confidence interval is bounded by 0.111 and 0.305 and excludes zero [35]. The upward movement of the point estimate is consistent with classical attenuation that arises from measurement error in self reported network data. The core conclusions are therefore stable and endogeneity does not substantially interfere with the empirical results.

5. Research Conclusions

We use survey data on 2,318 young workers in six Chinese cities, combined with a reconstructed job referral network of 9,742 weighted ties, to test the mechanism linking social capital to employment quality. The results show that social capital has a significant positive effect on the employment quality of Chinese youth. Resource access measured through the position generator and structural openness measured through betweenness and aggregate constraint are converted into contractual security, working time quality and development prospects, not merely into higher earnings. The main effect hypothesis is therefore verified.

At the level of the transmission mechanism, job search information breadth partially mediates the association between the two constructs. Social capital directly raises employment quality and further contributes to it by enlarging the set of distinct vacancy signals from which a match is selected. The indirect path accounts for roughly 30 percent of the total effect. The informational channel is thus substantial without exhausting the influence of network resources, and the residual share is attributable to the signalling channel through which a referral resolves screening uncertainty on the employer side. The estimation results support the complete transmission path.

The heterogeneity tests show that institutional status and educational attainment generate significant differentiation in the above effects. The value of network resources is larger among holders of rural household registration than among holders of urban registration, and larger among young workers below the tertiary threshold than among tertiary qualified workers. Social capital therefore operates as a partial substitute for institutional access and for formal credentials. Policies restricting informal recruitment channels without simultaneously improving formal matching infrastructure would consequently fall most heavily on the groups already least protected. The exponential family random graph estimates further show that referral ties form preferentially between individuals of similar occupational status. The network structure itself therefore reproduces the stratification of informational access across the youth labour market.

Robustness tests conducted by replacing the dependent variable with the deflated hourly wage and by instrumenting social capital with the leave one out mean of peers in the same neighbourhood and birth cohort indicate that the core conclusions are stable. Three limitations should be noted. The cross sectional design restricts the interpretation of the estimates to associations that survive an instrumental variable correction and not to fully identified causal effects. The referral network is reconstructed within the sampled population and therefore truncates ties reaching outside the six survey sites. The measurement of information breadth relies on retrospective self reports and is subject to recall error. Panel collection across successive labour market transitions and the linkage of survey responses to platform level application records would address these limitations and allow the estimation of how network resources are accumulated and depleted over the early career.

References

[1] International Labour Organization. (2024). Global employment trends for youth 2024: Decent work, brighter futures. International Labour Office.

[2] Dasgupta, S. (2022). Global employment trends for youth 2022: Investing in transforming futures for young people. International Labour Office.

[3] International Labour Organization. (2025). World employment and social outlook: Trends 2025. International Labour Office.

[4] OECD. (2021). OECD employment outlook 2021: Navigating the COVID-19 crisis and recovery. OECD Publishing.

[5] OECD. (2023). OECD employment outlook 2023: Artificial intelligence and the labour market. OECD Publishing.

[6] Eurofound. (2021). Working conditions and sustainable work: An analysis using the job quality framework. Publications Office of the European Union.

[7] Eurofound. (2022). Working conditions in the time of COVID-19: Implications for the future. Publications Office of the European Union.

[8] Warhurst, C., & Knox, A. (2022). Manifesto for a new quality of working life. Human Relations, 75(2), 304-321. https://doi.org/10.1177/00187267211027020

[9] Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., et al. (2022). Social capital I: Measurement and associations with economic mobility. Nature, 608(7921), 108-121. https://doi.org/10.1038/s41586-022-04996-4

[10] Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., et al. (2022). Social capital II: Determinants of economic connectedness. Nature, 608(7921), 122-134. https://doi.org/10.1038/s41586-022-04997-3

[11] Rajkumar, K., Saint-Jacques, G., Bojinov, I., Brynjolfsson, E., & Aral, S. (2022). A causal test of the strength of weak ties. Science, 377(6612), 1304-1310. https://doi.org/10.1126/science.abm7530

[12] Wheeler, L., Garlick, R., Johnson, E., Shaw, P., & Gargano, M. (2022). LinkedIn(to) job opportunities: Experimental evidence from job readiness training. American Economic Journal: Applied Economics, 14(2), 101-125. https://doi.org/10.1257/app.20200408

[13] Barwick, P. J., Liu, Y., Patacchini, E., & Wu, Q. (2023). Information, mobile communication, and referral effects. American Economic Review, 113(5), 1170-1207. https://doi.org/10.1257/aer.20211740

[14] Abebe, G., Caria, A. S., Fafchamps, M., Falco, P., Franklin, S., & Quinn, S. (2021). Anonymity or distance? Job search and labour market exclusion in a growing African city. The Review of Economic Studies, 88(3), 1279-1310. https://doi.org/10.1093/restud/rdaa072

[15] Glitz, A., & Vejlin, R. (2021). Learning through coworker referrals. Review of Economic Dynamics, 42, 37-71. https://doi.org/10.1016/j.red.2021.01.003

[16] Kircher, P. (2022). Schumpeter lecture 2022: Job search in the 21st century. Journal of the European Economic Association, 20(6), 2317-2352. https://doi.org/10.1093/jeea/jvac049

[17] Small, M. L., Perry, B. L., Pescosolido, B., & Smith, E. B. (2021). Personal networks: Classic readings and new directions in egocentric analysis. Cambridge University Press.

[18] Burt, R. S. (2021). Structural holes capstone, cautions, and enthusiasms. In M. L. Small, B. L. Perry, B. Pescosolido, & E. B. Smith (Eds.), Personal networks: Classic readings and new directions in egocentric analysis (pp. 384-416). Cambridge University Press.

[19] Borgatti, S. P., Everett, M. G., Johnson, J. C., & Agneessens, F. (2022). Analyzing social networks using R. SAGE Publications Ltd.

[20] Borgatti, S. P., Everett, M. G., Johnson, J. C., & Agneessens, F. (2024). Analyzing social networks (3rd ed.). SAGE Publications Ltd.

[21] Moro, E., Frank, M. R., Pentland, A., Rutherford, A., Cebrian, M., & Rahwan, I. (2021). Universal resilience patterns in labor markets. Nature Communications, 12(1), 1972. https://doi.org/10.1038/s41467-021-22223-z

[22] del Rio-Chanona, R. M., Mealy, P., Beguerisse-Díaz, M., Lafond, F., & Farmer, J. D. (2021). Occupational mobility and automation: A data-driven network model. Journal of the Royal Society Interface, 18(174), 20200898. https://doi.org/10.1098/rsif.2020.0898

[23] Lehdonvirta, V. (2022). Cloud empires: How digital platforms are overtaking the state and how we can regain control. The MIT Press.

[24] Coleman, J. S. (1988). Social capital in the creation of human capital. American Journal of Sociology, 94(Supplement), S95-S120. https://doi.org/10.1086/228943

[25] Lin, N. (2001). Social capital: A theory of social structure and action. Cambridge University Press.

[26] Lin, N., & Dumin, M. (1986). Access to occupations through social ties. Social Networks, 8(4), 365-385. https://doi.org/10.1016/0378-8733(86)90003-1

[27] Granovetter, M. S. (1973). The strength of weak ties. American Journal of Sociology, 78(6), 1360-1380. https://doi.org/10.1086/225469

[28] Burt, R. S. (1992). Structural holes: The social structure of competition. Harvard University Press.

[29] Soda, G., Mannucci, P. V., & Burt, R. S. (2021). Networks, creativity, and time: Staying creative through brokerage and network rejuvenation. Academy of Management Journal, 64(4), 1164-1190. https://doi.org/10.5465/amj.2019.1169

[30] Peixoto, T. P. (2023). Descriptive vs. inferential community detection in networks: Pitfalls, myths and half-truths. Cambridge University Press.

[31] Krivitsky, P. N., Hunter, D. R., Morris, M., & Klumb, C. (2023). ergm 4: New features for analyzing exponential-family random graph models. Journal of Statistical Software, 105(6), 1-44. https://doi.org/10.18637/jss.v105.i06

[32] Zhu, Y., Tian, D., & Yan, F. (2020). Effectiveness of entropy weight method in decision-making. Mathematical Problems in Engineering, 2020, 1-5. https://doi.org/10.1155/2020/5264537

[33] Hansen, B. E. (2022). Econometrics. Princeton University Press.

[34] Huber, M. (2023). Causal analysis: Impact evaluation and causal machine learning with applications in R. The MIT Press.

[35] Lee, D. S., McCrary, J., Moreira, M. J., & Porter, J. (2022). Valid t-ratio inference for IV. American Economic Review, 112(10), 3260-3290. https://doi.org/10.1257/aer.20210733