Dynamics of Educational Mismatch among Indigenous Workers in Peru

HECTOR BOTELLO

[hector.botello@wsu.edu]

November 24, 2025

Abstract

While the wage penalties associated with overeducation are well-documented in development economics, the *temporal dynamics* of this phenomenon remain understudied. This paper tests whether educational mismatch represents a transitory search friction (Career Mobility Theory) or a structural trap (Segmentation Theory) for indigenous populations in Peru. Exploiting the ENAHO Longitudinal Panel (2015–2019), we estimate a Cox proportional hazards model to analyze the duration of overeducation spells. Results reveal a significant structural "ethnicity penalty": indigenous workers are approximately 11% less likely to exit mismatched employment than their non-indigenous peers, conditional on human capital and spatial sorting. We document two mechanisms driving this persistence: a "double jeopardy" effect in rural labour markets and a severe "scarring" effect where the penalty widens significantly over the life cycle. These findings reject the hypothesis of mismatch as a stepping stone for indigenous graduates, supporting instead a model of labour market segmentation where credential devaluation functions as a glass floor.

Keywords: Overeducation; Indigenous Inequality; Duration Analysis; Labour Market Segmentation; Peru.
JEL Codes: J15, J24, O15, C41.

1 Introduction

For decades, the implicit social contract in Latin America—and particularly in the Andean region—has been remarkably simple: accumulate human capital, and you will escape the gravitational pull of poverty. This narrative drove a massive expansion of tertiary education in Peru, fueled by both public investment and a deregulation of private universities that promised a path to the middle class for historically marginalized groups.

While this massification is often viewed as a victory for development, the labor market complicates this optimistic narrative. When the supply of credentialed workers outpaces the creation of high-skilled jobs, we encounter the friction of *education-job mismatch*. While this inefficiency affects the aggregate economy, there is a growing suspicion that the burden settles disproportionately on those at the periphery of the labor market: the indigenous populations.

Ideally, mismatch acts as a temporary search friction—a meritocratic "stepping stone." However, Peruvian data suggests a structural trap. Despite high educational attainment, Quechua, Aymara, and Amazonian graduates face systematic credential devaluation. This segmentation undermines human capital as a vehicle for social mobility, effectively replacing the stepping stone with a discriminatory "glass floor."

Previous scholarship has struggled to capture the mechanics of this exclusion. While a substantial body of literature—anchored by Hall and Patrinos (2006) and Ñopo (2012)—documents ethnic wage gaps using Oaxaca-Blinder decompositions, these studies rely predominantly on cross-sectional data. Such static frameworks identify the *existence* of a penalty but obscure its *mechanism*. By failing to observe temporal dynamics, existing literature conflates temporary search frictions with structural exclusion, making it impossible to distinguish between transient underemployment and a permanent low-mobility trap.

This is the specific gap this study addresses. We pivot from static wage analysis to dynamic duration analysis to measure exactly how long this penalty lasts. Utilizing the **ENAHO Longitudinal Panel (2015–2019)**, we track individuals over time to test two competing theories: Sicherman (1991)'s "Career Mobility Theory," which views overeducation as a temporary phase, versus Thurow (1975)'s "Job Competition Model," which implies a queuing system where social hierarchies trap certain groups at the bottom. We build on Clark et al. (2017)'s analysis of duration in the U.S., adapting the framework to the stratified reality of the Peruvian labor market to analyze the intersection of ethnicity and the *persistence* of overeducation.

The primary objective of this study is to estimate the "ethnicity penalty" in the duration of mismatch. We test the hypothesis that indigenous workers exhibit a significantly lower hazard rate—a lower probability of exiting overeducation into a matched job—compared to non-indigenous workers, even after controlling for migration, education quality, and informal sector dynamics.

The stakes of this distinction are high. If we misdiagnose a structural trap as a temporary transition, policy interventions will remain toothless. Persistent mismatch signals to rural communities that investing in higher education yields diminishing returns, entrenching intergenerational poverty not because the poor lack skills, but because the market refuses to recognize them. By identifying whether indigenous graduates are trapped, we challenge the universality of human capital models in segmented economies and suggest

that Peru's "education for all" strategy requires a serious rethink regarding demand-side policies.

The remainder of this paper is organized as follows. Section 2 reviews the theoretical tension between mobility and segmentation. Section 3 describes the ENAHO panel data and the construction of mismatch variables. Section 4 details the empirical strategy, specifically the use of Cox proportional hazard models to handle right-censored duration data. Finally, we present results quantifying the persistence of the trap, followed by a discussion on policy implications for inclusive growth in the Andean region.

2 Theoretical Mechanisms of Mismatch

Two competing theoretical frameworks dominate the literature on overeducation, offering divergent predictions regarding its persistence.

Sicherman (1991) posits that overeducation is a temporary disequilibrium. In this view, workers voluntarily accept positions for which they are overqualified to acquire firm-specific human capital or simply to enter the labor market. Consequently, overeducation serves as a "stepping stone" toward better matches. If this theory holds, empirical analysis should observe high hazard rates (exit rates) from mismatch into matched employment, regardless of ethnicity, once workers demonstrate their productivity.

Conversely, Thurow (1975) argues that productivity resides in the job rather than the worker. When the supply of educated labor exceeds demand, a queue forms where education serves primarily as a screening device. In this "Job Competition" model, rigidities prevent wages from clearing the market, and surplus education becomes a defensive necessity. For marginalized groups, Arrow (1973) suggests that statistical discrimination leads employers to discount the credentials of specific demographic groups based on perceived average productivity. This framework predicts that for indigenous workers, overeducation is not a strategic choice but a 'waiting room' with no exit. Figure **??** illustrates these divergent theoretical trajectories While early cross-sectional studies treated mismatch as a static state, recent longitudinal analyses challenge the mobility hypothesis. Clark et al. (2017) utilized the National Longitudinal Survey of Youth (NLSY79) to demonstrate that overeducation is persistent in the United States, with past history of mismatch significantly predicting future mismatch. Similarly, Mavromaras et al. (2013) found that in Australia, state dependence is strong: workers trapped in mismatch tend to remain there, particularly those with lower unobserved ability.

However, these studies focus predominantly on developed economies. Evidence from developing contexts is sparse. Botello-Peñaloza (2015) and Quinn and Rubb (2006) analyzed Mexico and Colombia, finding that labor market rigidities exacerbate mismatch. Yet, these studies rely mostly on cross-sectional data, which prevents the identification of duration dependence and the separation of unobserved heterogeneity from true state

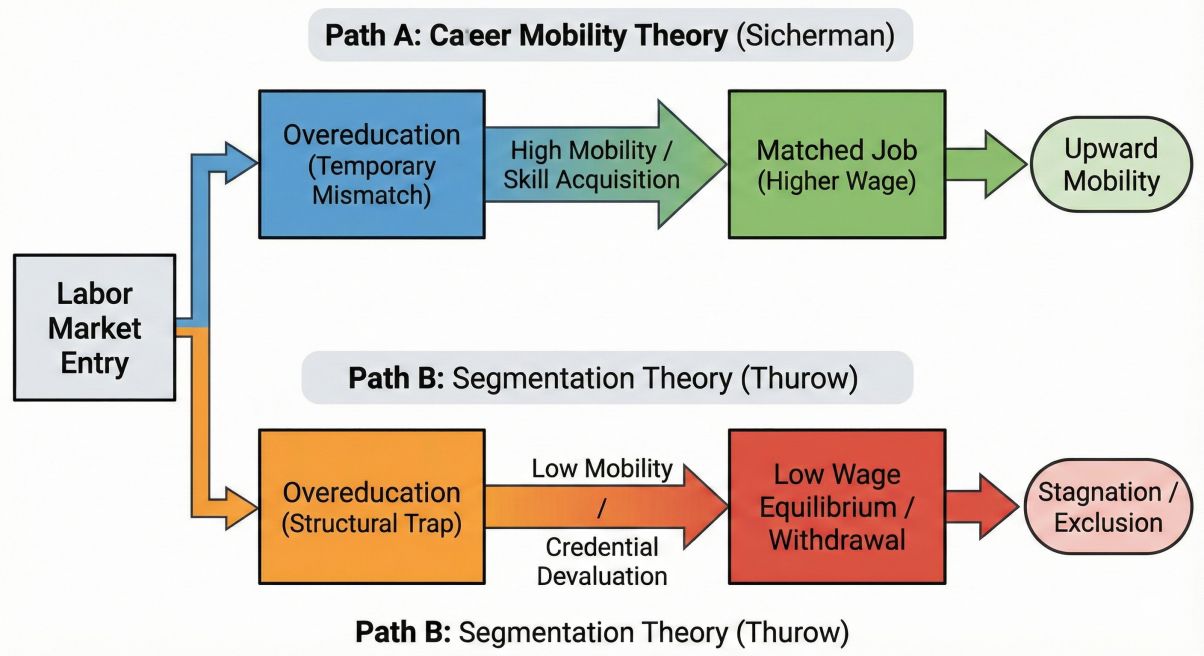


Figure 1: Conceptual Framework: Mobility vs. Segmentation Trajectories

dependence.

The literature on ethnic inequality in Latin America consistently identifies a severe penalty for indigenous populations. Ñopo (2012) and Hall and Patrinos (2006) document that even after controlling for human capital characteristics (education, age, experience), a significant "unexplained" wage gap remains for indigenous workers in Peru, Bolivia, and Ecuador. Ñopo (2008) attributes a portion of this gap to occupational segregation—indigenous workers are crowded into low-productivity sectors.

Few studies explicitly connect labor market segmentation to the persistence of mismatch. Castro and Yamada (2016), for example, analyzes static returns to education but fails to test for dynamic improvements in match quality. By treating degree acquisition as a sufficient condition for integration, the literature ignores differential duration dependence—specifically, the structural retention of indigenous workers in overeducated states.

A recurrent limitation in the reviewed literature is the failure to address unobserved heterogeneity in duration models. As noted by Heckman and Singer (1984), failure to control for unobserved characteristics (such as motivation or innate ability) can create spurious negative duration dependence. A worker might appear "trapped" in overeducation not because of structural barriers, but because they possess lower unobserved skill levels.

Most studies on indigenous inequality in Peru use Oaxaca-Blinder decompositions on repeated cross-sections (e.g., ENAHO annual data). This approach measures the wage gap at a specific point in time but cannot observe the trajectory of individual workers.

Without leveraging the panel dimension of datasets like the ENAHO (2015-2019), it is impossible to determine if the ethnic penalty is a result of lower entry wages or lower mobility rates.

In a nutshell, the literature presents a Human Capital Theory predicts that increased education should narrow ethnic wage gaps. However, empirical evidence shows that gaps persist. The missing link is the *dynamic matching process*.

1. We know overeducation hurts wages (Verdugo and Verdugo, 1989).
2. We know indigenous workers face discrimination (Ñopo, 2012).
3. We do not know if indigenous workers face a "duration penalty" that is, if they are significantly slower to exit mismatched jobs than non-indigenous workers with identical credentials.

This study aims to fill this specific gap. By applying survival analysis to the ENAHO Panel, this research will test whether the "Stepping Stone" hypothesis applies equally across ethnic groups or if the labor market is segmented by ethnicity, creating a "trap" for indigenous human capital. This distinction is vital for policy; if the issue is a trap, supply-side policies (more education) will yield diminishing returns without demand-side interventions (anti-discrimination enforcement).

3 Methodology and Empirical Strategy

3.1 Data

This study utilizes the *Encuesta Nacional de Hogares* (ENAHO) for the period 2015–2019, administered by the *Instituto Nacional de Estadística e Informática* (INEI). The ENAHO is the primary source of household information in Peru, employing a probabilistic, stratified, multi-stage, and independent sampling design. The survey coverage is national, encompassing both urban and rural areas across all 24 departments and the Constitutional Province of Callao.

The longitudinal sample is derived from a subset of the master sample. Specifically, the 2015-2019 panel comprises approximately 9,152 dwellings selected from 5 sub-samples of the primary sampling frame. While the strictly balanced panel (households interviewed in all five consecutive years) consists of approximately 1,866 households, restricting the analysis to this group would introduce significant survivorship bias, particularly among high-mobility indigenous populations. Consequently, we utilize an unbalanced panel approach, retaining individuals who appear in at least two waves. This maximizes the sample size and statistical power, allowing for a better analysis of transition dynamics.

The sample is restricted to the active labor force aged 14 to 65 who have completed at least primary education. To ensure national representativeness and correct for complex survey design, we utilize the baseline population weights (derived from the 2007 Census projections). For the survival analysis, the weight assigned to an individual in their first observed year is treated as time-invariant to ensure consistent estimation of the population parameters.

3.1.1 Variable Construction

1. **Ethnicity (ETH_i):** A binary variable constructed from self-identification. It takes the value of 1 if the individual identifies as Quechua, Aymara, or Native Amazonian, and 0 if they identify as Mestizo or White.

2. **Educational Mismatch (M_{it}):** We employ the *Realized Matches* (statistical) method, a widely accepted approach in labor economics (Verdugo and Verdugo, 1989; Groot and Maassen van den Brink, 2000). This method relies on the distribution of education within each occupation to determine the required level of schooling. Let S_{it} be the years of schooling of individual i at time t. For each 2-digit ISCO occupation j, we calculate the mean (μ_{jt}) and standard deviation (σ_{jt}) of schooling years. An individual is classified as Overeducated if $S_{it} > \mu_{jt} + \sigma_{jt}$.

3. **Duration of Mismatch (T_i):** We define a "spell" of overeducation as a continuous period during which an individual is classified as overeducated. The duration is measured in years.

3.2 Econometric Model

To test the hypothesis that indigenous workers face a "duration penalty" in escaping educational mismatch, we employ survival analysis techniques. We define the "failure event" (exit) as the transition from an overeducated state to a matched state. Observations that remain mismatched at the end of the panel or exit the sample for other reasons are treated as right-censored.

We estimate a semi-parametric Cox Proportional Hazards model (Cox, 1972). This approach allows us to estimate the effect of ethnicity on the hazard rate of exiting overeducation without imposing a restrictive functional form on the baseline hazard. The hazard function for individual i at time t is specified as:

$$h(t|\mathbf{X}_{it}) = h_0(t) \exp\left(\beta_1 ETH_i + \beta_2 S_{it} + \gamma \mathbf{Z}_{it}\right) \tag{1}$$

Where:

- $h(t|\mathbf{X}_{it})$ is the hazard rate, representing the probability that an individual exits overeducation at time t, conditional on remaining overeducated up to time t.
- $h_0(t)$ is the non-parametric baseline hazard.
- ETH_i is the time-invariant ethnicity dummy. A hazard ratio ($\exp(\beta_1)$) less than 1 indicates a lower probability of exit, implying a "trap" or penalty, supporting the segmentation hypothesis (Thurow, 1975).
- S_{it} represents human capital variables, specifically years of schooling and potential experience (approximated as $Age - Schooling - 6$).
- $\mathbf{Z}_{it}$ is a vector of control variables designed to isolate the ethnicity effect from other structural factors.

We estimate three nested models to observe how the ethnicity coefficient changes with the inclusion of controls:

- **Model 1 (Unconditional):** Includes only the ethnicity dummy to estimate the raw differential.
- **Model 2 (Full Specification):** Adds controls for years of schooling, gender, and age groups (14-24, 25-34, 35-44, 45-54, 55+). This tests whether the duration gap is driven by differences in education or life-cycle stages (Sicherman, 1991). Include labor market and geographical controls, a dummy for Rural area and fixed effects for macro-regions (Coast, Sierra, Selva, Lima). Also include an indicator for Informal employment and categorical variables for Firm Size (Micro, Small, Medium, Large).

All models are estimated using standard errors clustered at the individual level to account for serial correlation within individuals across the panel. The Proportional Hazards assumption is verified using Schoenfeld residuals.

To investigate whether the ethnicity penalty is exacerbated by spatial or demographic factors, we extend the Cox model to include interaction terms. We test the interaction between ethnicity and rural location ($ETH_i \times Rural_{it}$) to determine if the penalty is distinct from the general disadvantage of rural labor markets.

3.3 Robustness Strategy

To ensure that the estimated ethnicity penalty is not an artifact of unobserved structural factors or model misspecification, we implement several robutness test. the first one is a falsification (placebo) test.

We restrict the analytical sample exclusively to non-indigenous workers and randomly assign a "placebo" indigenous status ($Placebo_i$) using a uniform distribution, such that

$P(Placebo_i = 1) = 0.5$. We then re-estimate the fully specified Cox proportional hazards model (Equation 1) using this synthetic regressor in place of the true ethnicity variable.

Since $Placebo_i$ is orthogonal to individual productivity and labor market characteristics by design, we expect the estimated hazard ratio to be statistically indistinguishable from unity. A significant coefficient in this specification would imply that our model is capturing spurious duration dependence or background noise rather than a true structural penalty associated with indigenous identity.

3.3.1 Addressing Selection Bias via Competing Risks

A critical limitation of the standard Cox proportional hazards model is the assumption of *independent censoring*. This assumption implies that individuals who exit the risk set for reasons other than the primary failure event (finding a matched job) do so randomly, and that their probability of finding a match in the future remains identical to those who remain in the sample.

In the context of ethnic inequality in Peru, this assumption is likely violated. Indigenous workers may face higher barriers to employment, leading to a higher propensity for labor force withdrawal or transitions into unemployment—phenomena widely documented as the "discouraged worker effect." If indigenous workers are systematically more likely to exit the sample into inactivity ($T_{inactivity} < T_{match}$), a standard Cox model treats these exits as censored observations, effectively assuming they are still searching for employment with the same success rate as active searchers. This would artificially inflate the estimated hazard rate for indigenous workers, masking the true extent of the "ethnicity penalty."

To address this potential bias, we estimate a Competing Risks Regression model based on the proportional subdistribution hazards method proposed by Fine and Gray (1999). Unlike the standard Cox model, which estimates the cause-specific hazard, the Fine and Gray model estimates the subdistribution hazard, $h_1(t)$, defined as:

$$h_1(t|\mathbf{X}) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t, \text{Event} = 1 | T \geq t \cup (T \leq t \cap \text{Event} \neq 1))}{\Delta t} \tag{2}$$

This specification explicitly acknowledges that individuals who experience the competing event (unemployment or inactivity) remain in the risk set for the event of interest (finding a match) but with a probability of zero. By accounting for the informative nature of exits to unemployment, this approach provides a robust estimate of the structural disadvantage faced by indigenous workers, conditional on their ability to remain in the active labor force.

Table 1: Descriptive Statistics of Overeducated Workers by Ethnicity

	Non-Indigenous		Indigenous	
Variable	Mean	SD	Mean	SD
Years of Schooling	14.29	1.94	14.00	1.98
Monthly Income (S/.)	1240.60	1457.69	907.05	977.13
Age	36.48	12.81	34.23	12.03
Informality Indicator	1.48	0.50	1.34	0.47
Rural Resident	0.095	0.29	0.21	0.41
Duration of Spell (Years)	1.17	0.49	1.20	0.56

Source: Author's calculations using ENAHO Panel 2015-2019.

4 Results

This section presents the empirical findings regarding the duration of educational mismatch among indigenous workers in Peru. We begin with descriptive statistics of the risk set—workers who experience overeducation—followed by non-parametric survival estimates and, finally, the results from the Cox proportional hazards models.

4.1 Descriptive Statistics of the Mismatched Population

Table 1 summarizes the characteristics of workers who experienced at least one spell of overeducation during the 2015–2019 period. The data reveals disparities between indigenous and non-indigenous workers within this mismatched group.

First, despite being similarly overeducated, indigenous workers face a wage penalty. The average monthly income for overeducated indigenous workers is 907 soles, approximately 27% less than the 1,241 soles earned by their mestizo/white counterparts. This confirms that the "ethnicity penalty" manifests immediately in wages, even before considering the duration of the mismatch.

Second, while the average years of schooling are slightly lower for indigenous workers (14.0 years vs. 14.3 years), both groups possess education levels well above the national average, consistent with the definition of overeducation. Notably, indigenous workers in this sample are more concentrated in rural areas (21% vs. 9.5%), highlighting the intersection of ethnicity and geography.

Finally, the raw mean duration of overeducation spells is slightly higher for indigenous workers (1.20 years) compared to non-indigenous workers (1.17 years). While this difference appears small in raw means, it is important to note that these figures are right-censored; many spells of overeducation were ongoing at the end of the survey period. Survival analysis is thus required to properly estimate the differential persistence of these spells.

4.2 Survival Analysis: The Persistence of Mismatch

To formally test the "Ethnicity Penalty" hypothesis, we estimate the probability of exiting overeducation using the Cox Proportional Hazards model. The results are presented in Table 2, reporting Hazard Ratios (HR). An HR below 1 indicates a lower probability of exiting mismatch (i.e., longer persistence), while an HR above 1 indicates faster exit.

Model 1 (Unconditional) shows a hazard ratio of 0.91 ($p = 0.077$) for indigenous workers. This suggests that, without controlling for any other factors, indigenous workers are approximately 9% less likely to exit a mismatched job in any given year compared to non-indigenous workers. This provides initial evidence of a duration penalty.

Model 2 (Full Specification) Once we control for the fact that indigenous workers are overrepresented in rural areas—where exit rates are artificially high due to labor force withdrawal or subsistence work—the penalty is still negative. The hazard ratio for indigenous workers drops to 0.89 ($p = 0.067$).

The central finding, after adjusting for human capital and geographic location, indigenous workers face an "ethnicity penalty" of approximately 11% in their rate of exiting educational mismatch. They are significantly more likely to remain trapped in overeducation than a non-indigenous worker with identical skills living in the same region.

Table 2: Cox Proportional Hazards Models: Probability of Exiting Overeducation

	(1) Unconditional	(2) Full Specification
Indigenous	0.911*	0.894*
	(0.048)	(0.055)
Years of Schooling		1.226***
		(0.008)
Female		0.838***
		(0.041)
Rural		1.516***
		(0.086)
Observations	340,676	235,701

Coefficients are hazard ratios. Standard errors clustered by individual in parentheses.
* p<0.1, ** p<0.05, *** p<0.01. The second model includes regional and firm-size fixed effects.

4.3 Heterogeneity Analysis

The aggregate results in Table 2 established an average ethnicity penalty. However, to test the mechanisms of this disadvantage, we extend the Cox model to examine two classes of heterogeneity: spatial segmentation and life-cycle dynamics. The results are presented in Table 3.

Column (1) tests the interaction between indigenous identity and rural residence ($Indigenous \times Rural$). The main effect for indigenous workers (representing the urban indigenous population) is 0.90 ($p<0.10$). This indicates that even in urban centers like Lima or Arequipa—where labor markets are deeper and more diverse—indigenous workers still face a 10% lower probability of exiting overeducation compared to their mestizo neighbors. The penalty is not merely a artifact of rural underdevelopment; it is structural.

Second, the interaction term is 0.75 ($p<0.01$) and significant. This suggests a "double jeopardy" for indigenous workers in rural areas. While living in a rural area reduces mobility for everyone (Rural HR = 0.90), being indigenous in a rural area reduces the hazard rate by an additional multiplicative factor. This show that rural labor markets stratify by ethnicity, potentially due to stronger occupational segregation in traditional agrarian economies.

Column (2) explores whether the trap tightens over time by interacting ethnicity with age groups. The reference group is the youngest cohort (14–24 years).

A striking pattern emerges: the ethnicity penalty is virtually nonexistent for young entrants. The coefficient for indigenous workers (young cohort less than 25 years old) is 1.10 and statistically insignificant, suggesting that young indigenous graduates exit mismatch at rates comparable to their peers. However, the interaction terms reveal a severe divergence as workers age.

For indigenous workers aged 25–34, the hazard ratio drops significantly relative to the base ($HR_{int} = 0.76$). This decline accelerates in older cohorts, reaching a hazard interaction of 0.49 ($p<0.01$) for those over 55. This finding supports a "cumulative disadvantage" or "scarring" hypothesis. While indigenous workers may enter the labor market with similar mobility potential, those who fail to find matches early find themselves in a trap that hardens significantly over time, whereas non-indigenous workers sustain higher mobility rates into their mid-career.

4.4 Robustness Check

To ensure that the estimated ethnicity penalty is not an artifact of unobserved structural factors affecting the general population, we conduct a falsification (placebo) test. We restrict the sample exclusively to non-indigenous workers ($N = 267,388$) and randomly assign a "placebo" indigenous status to a subset of individuals using a random uniform distribution. If the penalty identified in our main specification were driven by model misspecification or spurious duration dependence, we would expect to observe a significant coefficient for this false treatment.

The results, presented in Table 4, confirm the validity of our identification strategy. The estimated hazard ratio for the placebo group is 0.97 ($p = 0.494$), which is statis-

Table 3: Heterogeneity in Exit Rates: Geography and Life Cycle

	(1) Spatial Interaction	(2) Life Cycle Interaction
Main Effects		
Indigenous	0.903*	1.110
	(0.049)	(0.108)
Rural	0.896**	
	(0.046)	
Interactions		
Indigenous × Rural	0.749***	
	(0.062)	
Indigenous × Age 25–34		0.761**
		(0.093)
Indigenous × Age 35–44		0.666***
		(0.082)
Indigenous × Age 45–54		0.591***
		(0.077)
Indigenous × Age 55+		0.490***
		(0.069)
Controls		
Female	0.575***	0.569***
	(0.020)	(0.020)
Regional Fixed Effects	Yes	Yes
Age Group Dummies	No	Yes
Observations	379,597	379,596

Coefficients reported are Hazard Ratios. Standard errors clustered by individual in parentheses. * p<0.1, ** p<0.05, *** p<0.01. Reference Age Group: 14–24.

Table 4: Placebo Test: False Treatment on Non-Indigenous Sample

	(1) Placebo Specification
Placebo Indigenous	0.973 (0.039) [p=0.494]
Controls	
Years of Schooling	1.238***
Female	0.650***
Rural	1.821***
Regional Fixed Effects	Yes
Age Group FE	Yes
Observations	267,388

Notes: Sample restricted to non-indigenous workers. "Placebo Indigenous" is a randomly assigned binary variable. Coefficients reported are Hazard Ratios. Standard errors clustered by individual in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

tically indistinguishable from unity. This indicates that "fake" indigenous workers exit overeducation at the same rate as the control group. Consequently, we can attribute the duration penalty observed in the main analysis specifically to the indigenous condition rather than to background noise in the data generating process.

5 Discussion

This study set out to determine whether the labor market experiences of indigenous workers in Peru support the Career Mobility Theory—where educational mismatch is a temporary stepping stone—or the Segmentation Theory, where it functions as a structural trap. By utilizing the ENAHO 2015–2019 longitudinal panel, we provide the first dynamic estimates of the "ethnicity penalty" in the duration of overeducation in the Andean region. Our results offer strong evidence in favor of labor market segmentation, revealing that indigenous workers face a persistent disadvantage in exiting mismatched employment, a finding that remains robust even after controlling for human capital and geographic heterogeneity.

5.1 The Structural Nature of the Ethnicity Penalty

Our primary finding is that indigenous workers are approximately 11% less likely to exit a state of overeducation in any given year compared to their non-indigenous peers

($HR \approx 0.89$, $p < 0.10$). This result challenges the universality of Sicherman (1991)'s Career Mobility Hypothesis in developing economies. While mobility theory posits that overeducated workers will exhibit high turnover as they search for better matches, our survival analysis indicates that for indigenous populations, this mobility is significantly constrained. This aligns with the segmentation arguments posited by Thurow (1975) and Ñopo (2012), suggesting that indigenous workers effectively queue for high-skilled jobs but face higher barriers to entry, leading to longer spells of underutilization.

The magnitude of this penalty is consistent with cross-sectional wage gap literature in Latin America, which typically finds unexplained residuals of 20–30% (Hall and Patrinos, 2006; Atal et al., 2009). However, our dynamic approach adds a new dimension, the penalty is not merely a one-time wage loss at entry but a cumulative loss of human capital.

5.2 The Geography of Mismatch: Resolving the Empirical Puzzle

Our heterogeneity analysis (Table 3) reveals that the ethnicity penalty is not uniform; rather, it is compounded by geography and age.

First, we find evidence of a "double jeopardy" in rural markets. While the penalty exists in urban centers ($HR \approx 0.90$), the interaction between indigenous identity and rural residence is significant and negative ($HR \approx 0.75$). This contradicts the notion that rural labor markets are egalitarian in their stagnation; instead, they appear to stratify sharply by ethnicity, possibly due to rigid occupational segregation in traditional agrarian economies where indigenous workers are crowded into low-mobility sectors.

Second, the life-cycle analysis uncovers a disturbing "scarring" effect. Young indigenous entrants (14–24) exhibit exit rates comparable to their mestizo peers. However, the penalty widens drastically with age, with the hazard ratio dropping to 0.49 for workers over 55. This supports Mavromaras et al. (2013)'s hypothesis of negative state dependence: while young indigenous graduates may initially compete on equal footing, those who fail to match early find themselves in a trap that hardens significantly over time, whereas non-indigenous workers sustain higher mobility rates into their mid-career.

These findings necessitate a refinement of human capital theory when applied to stratified societies. The standard Mincerian assumption that returns to education are primarily a function of individual productivity relies on a frictionless matching process. Our results suggest that for indigenous workers, the "signaling" value of a degree (Spence, 1973) may be discounted by employers due to information asymmetries regarding the quality of rural education or statistical discrimination (Arrow, 1973). Consequently, an indigenous graduate must effectively "pay" a higher tenure tax in low-skill employment to prove their worth.

Certain limitations confirm the need for caution. First, our measure of overeducation relies on the statistical method (mean + 1 SD). While robust and common in the literature, it may understate mismatch in economies where the entire workforce is becoming overqualified due to credential inflation. Second, while we control for informality, the ENAHO panel does not allow us to perfectly distinguish between exits to matched employment versus exits to unemployment or inactivity. Future research using administrative tax data could more precisely track these destination states to decompose the competing risks of exit.

6 Conclusion

This study investigated the longitudinal dynamics of educational mismatch among indigenous workers in Peru. By exploiting the 2015–2019 ENAHO panel, we provided the first dynamic estimates of the "ethnicity penalty" in the Andean region. Our findings reject the Career Mobility Hypothesis for indigenous populations, supporting instead a Segmentation Theory perspective where overeducation functions as a structural trap.

We document a ethnicity penalty of approximately 11% in the probability of exiting mismatch. This disadvantage is masked in naïve models by the high turnover of rural labor markets but becomes evident once spatial sorting is accounted for. Furthermore, we identify a "scarring" mechanism: the penalty is nonexistent for young entrants but severe for older cohorts, suggesting that the trap tightens over the life cycle.

These results have profound implications for development policy. The aggressive expansion of tertiary education in Peru has successfully raised attainment but has not been matched by inclusive absorptive capacity. To prevent education from becoming a "dead end" for indigenous youth, policy must pivot from supply-side interventions to demand-side structural reforms. This includes strengthening anti-discrimination enforcement in hiring, reducing information asymmetries through better labor market signaling systems, and fostering regional economic development to create high-skilled demand outside of Lima. Ultimately, closing the ethnic gap requires acknowledging that in a segmented economy, equal education does not guarantee equal opportunity.

References

Arrow, K. (1973). The theory of discrimination. *Discrimination in Labor Markets*, 3(10):3–33.

Atal, J. P., Ñopo, H., and Winder, N. (2009). New century, old disparities: Gender and ethnic wage gaps in latin america. Working Paper IDB-WP-109, Inter-American Development Bank.

Botello-Peñaloza, H. A. (2015). Evolución de las causas y efectos del desajuste educativo en el mercado laboral mexicano 1990-2010. *Económicas CUC*, 36(1):275–304.

Castro, J. F. and Yamada, G. (2016). Returns to higher education in peru: Trends and heterogeneity. *Journal of International Development*, 28. Placeholder for user provided citation.

Clark, B., Joubert, C., and Maurel, A. (2017). The career prospects of overeducated americans. *IZA Journal of Labor Economics*, 6(1).

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2):187–202.

Fine, J. P. and Gray, R. J. (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*, 94(446):496–509.

Groot, W. and Maassen van den Brink, H. (2000). Overeducation in the labor market: A meta-analysis. *Economics of Education Review*, 19(2):149–158.

Hall, G. and Patrinos, H. A. (2006). Indigenous peoples, poverty and human development in latin america. *Palgrave Macmillan*.

Heckman, J. J. and Singer, B. (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica*, 52(2):271–320.

Mavromaras, K., McGuinness, S., O'Leary, N., Sloane, P., and Wei, Z. (2013). Over-skilling dynamics and education pathways. *Economics of Education Review*, 37:619–628.

Ñopo, H. (2008). Matching as a tool to decompose wage gaps. *The Review of Economics and Statistics*, 90(2):290–299.

Ñopo, H. (2012). *New Century, Old Disparities: Gender and Ethnic Earnings Gaps in Latin America and the Caribbean*. Inter-American Development Bank/World Bank, Washington, DC.

Quinn, M. A. and Rubb, S. (2006). Educational mismatches in mexico: Additional evidence of the importance of labor market assignments? *Economics of Education Review*, 25(2):147–156.

Sicherman, N. (1991). Overeducation in the labor market. *Journal of Labor Economics*, 9(2):101–122.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3):355–374.

Thurow, L. C. (1975). *Generating Inequality*. Basic Books.

Verdugo, R. R. and Verdugo, N. T. (1989). The impact of surplus schooling on earnings: Some additional findings. *The Journal of Human Resources*, 24(4):629–643.