

RESUMO EXPANDIDO (PÚBLICO GERAL/NÃO ASSOCIADOS AKB) /
EXTENDED ABSTRACTS (NON AKB MEMBERS) - ÁREA 5. COMÉRCIO
INTERNACIONAL, TAXA DE CâMBIO E DESENVOLVIMENTO ECONÔMICO

**ECONOMIC COMPLEXITY, REGIONAL DEVELOPMENT, AND WAGE
INEQUALITY: AN ECONOMETRIC ANALYSIS OF BRAZILIAN LABOR
MARKETS**

Lucas Scheremetta Santos (scheremettalucas@gmail.com)

A prominent literature beginning with Hartmann et al. (2017) and Sbardella et al. (2017) reports a negative association between economic complexity, captured by the Economic Complexity Index (ECI), and income inequality across countries. This association has been interpreted through institutional channels, since more sophisticated productive structures are argued to foster inclusive institutions and to expand sectors that pay higher and more compressed wages. The finding has been influential in policy debates in Latin America, where productive sophistication is frequently advanced as a joint instrument for growth and equity. Whether this empirical regularity travels to subnational settings is an open question. Within a single country, key institutional variables that mediate the international relationship, most notably the legal minimum wage and labor market regulations, are held uniform across regions. Brazil offers a useful setting, since the federal minimum wage applies uniformly to a highly heterogeneous productive geography and the formal labor market is subject to nontrivial compression around the minimum (Engbom and Moser, 2022; Haanwinckel, 2023). The question sits at the intersection of the literatures on economic development, structural change, and regional labor markets. This

paper examines how regional economic complexity is associated with wage levels, wage inequality, and the structure of returns to worker characteristics across Brazilian Immediate Geographic Regions (IGRs).

The paper pursues three objectives. The first is to characterize the association between ECI and several moments of the regional wage distribution, including the Gini coefficient, percentile ratios, the Theil index decomposed into variation between and within sectors, and the logs of P10, P50 and P90, in order to identify where in the distribution the association is concentrated. The second is to estimate, at the worker level, whether the partial association between regional ECI and individual wages varies along the conditional wage distribution. The third is to decompose the wage gap between regions of high and low complexity into composition (differences in worker characteristics) and structural (differences in returns) components, using the recentered influence function decomposition in the Oaxaca-Blinder tradition.

The analysis uses administrative records at the individual level from the *Relação Anual de Informações Sociais* (RAIS) for 2024, covering formal employment contracts in Brazil. Standard filters are applied (active contracts on December 31, weekly hours of at least 30, ages 18 to 65, exclusion of Agriculture and Public Administration), and the resulting sample is nested in 510 IGRs. Originally proposed in the international trade literature to capture sophistication in countries' export baskets (Hidalgo and Hausmann, 2009), the ECI is constructed here from a matrix of employment by IGR and CNAE activity at the four digit level, using the binary specialization criterion (location quotient at or above one) and the eigenvector method. Regional controls for GDP per capita and population are obtained from IBGE, sectoral shares are computed from RAIS.

Three complementary empirical strategies are deployed. The first model estimates OLS regressions in a single cross section at the IGR level for eight dependent variables (Gini, Theil and its components, P90/P10, the variance of log wages, and the logs of P10 and P50), in linear and quadratic specifications in ECI, controlling for the share of workers with tertiary education, the share of industrial employment, log population, log GDP per capita and macroregion

dummies. The second model estimates conditional quantile regressions at the worker level for t in $\{0.10, 0.25, 0.50, 0.75, 0.90, 0.99\}$, with regional ECI entering alongside individual controls (years of schooling, age and age squared, sex, race or color, tenure, log establishment size, state fixed effects). The third model implements the recentered influence function decomposition between IGRs above and below the median ECI, at five quantiles and for the Gini coefficient, with standard errors obtained via 200 bootstrap replications. The empirical strategy is associational. Robustness exercises include standard errors clustered at the state level, exclusion of the largest IGRs, macroregion interactions, alternative ECI specifications (Mealy et al., 2019), coefficient stability across nested models, spatial autocorrelation diagnostics, and quantile estimation by complexity group.

Eight findings emerge. First, conditional on controls, ECI shows no detectable association with the regional P10 ($\beta = +0.003$, not significant), while the standard deviation of the P10 across IGRs (R\$ 0.41) is an order of magnitude smaller than that of the P90 (R\$ 7.69). About 91% of IGRs have a P10 at or below one minimum wage per hour, and all fall below 1.2 minimum wages per hour. The Spearman correlation between ECI and the share of workers earning at or below the minimum is -0.59. The pattern is consistent with the regional P10 being anchored by the federal minimum wage, with ECI relating to the share of workers near that level rather than to the level itself. Second, ECI is positively associated with the log of P50 ($\beta = +0.013$, $p = 0.038$) and negatively associated with measures of dispersion between the tails (P90/P10: $\beta = -0.157$, $p = 0.032$; variance of log wages: $\beta = -0.015$, $p = 0.011$), while the Gini coefficient itself is statistically indistinguishable from zero in conditional specifications. Quadratic specifications indicate nonlinearities in P10 and P50 ($p < 0.01$ for ECI squared). Third, in the conditional quantile regression at the worker level, the ECI coefficient increases monotonically across quantiles, from $\beta = 0.004$ at $t = 0.10$ to $\beta = 0.059$ at $t = 0.99$. Strict monotonicity holds for all adjacent quantile pairs, and the interquantile Wald test of equality between $t = 0.10$ and $t = 0.90$ is rejected at $p < 0.001$. The estimated partial association between regional ECI and individual wages is substantially larger for workers in upper positions of the conditional wage distribution. Fourth, the bootstrap implementation of the recentered influence function decomposition attributes between 52% and 66% of the wage gap between regions of high and low complexity to the structural component, with the remainder attributed to

composition. The structural component grows in absolute magnitude across quantiles, from 0.022 log points at $t = 0.10$ to 0.147 at $t = 0.90$, with tight bootstrap confidence intervals. Quantile regressions estimated separately by complexity group indicate that returns to schooling are systematically larger in regions of high complexity, with the gap widening from +0.024 at $t = 0.10$ to +0.036 at $t = 0.90$ per year of schooling.

Fifth, regions of high complexity display a share of the Theil index attributable to variation between sectors of 33%, against 44.5% in regions of low complexity ($t = -8.09$). The pattern is consistent with a relative redistribution of inequality from variation between sectors to variation within sectors as ECI rises. Sixth, interactions of ECI with macroregion dummies indicate that the positive association with the logs of P10 and P50 is concentrated in the Southeast (the base category), while in the North, Northeast, South and Central West the interaction terms are negative and significant ($p < 0.001$ for the log of P10 in the North, Northeast and Central West). The results suggest substantial heterogeneity in the relationship between ECI and wages within Brazil. Seventh, quantile estimates by complexity group indicate that penalties associated with female sex and with workers identifying as Black or Brown tend to be larger in regions of high complexity, with the differential growing at upper quantiles (the female penalty at $t = 0.90$ is -0.250 in regions of high complexity against -0.230 in regions of low complexity). Eighth, the bivariate coefficient on ECI in a Gini regression is +0.020; the inclusion of regional controls reverses the sign to -0.014, and the further addition of macroregion dummies brings it to -0.007 and statistical insignificance. Coefficient stability exercises suggest that bivariate ECI inequality associations substantially reflect covariation with human capital, GDP per capita and regional size.

The findings indicate that the negative association between ECI and inequality reported across countries does not extend in a straightforward manner to the Brazilian subnational case. Conditional on standard controls, ECI is not significantly associated with the Gini coefficient or with the regional P10, but is robustly associated with median wages, with reduced dispersion between the tails, with larger partial returns at the top of the conditional wage distribution, and with a relative shift of inequality from variation between sectors to variation within sectors. Decomposition results indicate that most of the wage gap

between regions of high and low complexity stems from differential returns to observable characteristics rather than from compositional differences. The findings carry implications for the development literature that views productive sophistication as a joint instrument for growth and equity, suggesting that within country distributive effects of complexity may operate through channels distinct from those at the cross country level. The paper is associational and based on a single cross section; the high spatial autocorrelation of residuals for the log of P50 (Moran's $I = 0.34$, $p < 0.001$) and the limited number of state clusters ($G = 27$) suggest that future versions should incorporate spatial econometric specifications, wild cluster bootstrap inference, and panel estimation with regional fixed effects.

Palavras-chave: economic complexity; regional development; wage inequality; quantile regression; rif oaxaca-blinder decomposition; minimum wage; labor market.