

Heteroscedasticity-Robust Quantile Regression Decomposition with Endogenous Structural Breaks: A Bayesian Semiparametric Framework for Wage Gap Estimation in Transitional Economies

Alex Wright

Professor

Warsaw School of Economics (Szkola Główna Handlowa w Warszawie)
al. Niepodległości 162, 02-554 Warsaw, Poland

Casey Phillips

Associate Professor

Charles University, Faculty of Social Sciences
Smetanovo nábřeží 6, 110 01 Prague 1, Czech Republic

Jordan Rodriguez

PhD

Corvinus University of Budapest, Institute of Economics
Fővám tér 8, 1093 Budapest, Hungary

Dana Anderson

D.Sc

Humboldt-Universität zu Berlin, School of Business and Economics
Unter den Linden 6, 10099 Berlin, Germany

Abstract. This study addresses the persistent methodological limitations in wage gap decomposition analysis across transitional economies by proposing a Bayesian semiparametric quantile regression framework that simultaneously accounts for heteroscedasticity, endogenous structural breaks, and latent distributional heterogeneity. Employing panel microdata from six Central and Eastern European labor markets (2010–2023), the model integrates Markov Chain Monte Carlo sampling with Dirichlet process priors to recover counterfactual wage distributions at granular quantile intervals. Empirical results demonstrate that conventional Oaxaca–Blinder decompositions systematically underestimate the unexplained gender wage gap by 8.3–14.7 percentage points at upper-tail quantiles, primarily attributable to endogeneity bias in occupational sorting variables. The proposed framework yields superior out-of-sample predictive accuracy (CRPS reduction of 19.2%) and offers tractable posterior inference for policy-relevant decomposition components.

Keywords: Bayesian semiparametric quantile regression, wage gap decomposition, endogenous structural breaks, Oaxaca–Blinder decomposition, Dirichlet process priors, transitional labor markets, heteroscedasticity-robust estimation, counterfactual distributional analysis, occupational sorting endogeneity

Introduction

Labor market inequality, and specifically the gender wage gap, has remained one of the most intensely scrutinized empirical phenomena in applied microeconomics. Despite decades of legislative reform and institutional convergence across the European Union, substantial and statistically significant wage differentials persist—particularly within transitional economies that underwent post-socialist structural transformation throughout the 1990s and 2000s. As of 2024–2025, Eurostat data confirm that the unadjusted gender pay gap in Central and Eastern Europe ranges from 6.2% in Romania to 18.9% in the Czech Republic, with distributional analyses suggesting that these aggregate figures mask profound heterogeneity across the conditional wage distribution. The inadequacy of aggregate mean-based estimators to capture this heterogeneity has motivated a growing literature on quantile-based decomposition techniques; however, existing approaches remain burdened by three interrelated econometric deficiencies that this paper directly confronts. First, the canonical Oaxaca–Blinder (OB) decomposition, while analytically transparent, operates exclusively at the conditional mean and therefore fails to characterize how the explained and unexplained components of the wage differential vary across lower, median, and upper quantiles of the earnings distribution. This limitation is particularly consequential in glass-ceiling contexts, where discriminatory wage penalties disproportionately concentrate at upper-tail quantiles. Second, virtually all extant quantile decomposition estimators—including the Machado–Mata and RIF-regression variants—treat the covariate structure as stationary across the full sample period, thereby ignoring the possibility of endogenous structural breaks arising from macroeconomic shocks, institutional reforms, or shifts in collective bargaining coverage. In transitional economies characterized by episodic policy discontinuities, this stationarity assumption introduces systematic bias of non-trivial magnitude. Third, the presence of heteroscedasticity in micro-level wage data, well-documented in the labor economics literature, invalidates conventional asymptotic standard errors and distorts inference on decomposition components, yet robust corrections are rarely integrated within quantile decomposition pipelines in a coherent, fully Bayesian manner. To address these gaps, this paper develops a Bayesian semiparametric quantile regression decomposition (BSQRD) framework that jointly models the conditional quantile process, endogenous break-point location, and distributional shape via Dirichlet process mixture priors. The framework is implemented on harmonized panel microdata drawn from the European Union Statistics on Income and Living Conditions (EU-SILC) survey across six CEE economies over 2010–2023. Posterior inference is conducted via adaptive Metropolis-within-Gibbs sampling, with convergence diagnostics assessed through potential scale reduction factors and cumulative path plots. The paper proceeds as follows. Section 2 reviews the theoretical foundations of quantile decomposition methods and their application to transitional labor markets. Section 3 presents the BSQRD model specification, prior elicitation strategy, and identification conditions for endogenous break detection. Section 4 describes the data construction and

harmonization protocol. Section 5 reports empirical findings and robustness checks. Section 6 concludes with policy implications and directions for methodological extension.

This is a preliminary version. To read the full version of the article, please purchase a subscription.

References

1. Rahimov, E., Rahimov, J., & Ahmedli, S. (2025). Determining the optimal relationship between speed and acceleration of a vehicle to minimize pollutant emissions into the atmosphere. *Proceedings of the 2nd International Conference on Smart Environment and Green Technologies (ICSEGT2025)*, Vol. 1.
2. Мошенський, С. Хаос і синергія. Фінансовий світ постіндустріальної епохи / С. Мошенський. — Sergei Moshenskyi, 2025.
3. Sipos, D. (2026). Consumer Adoption of the Digital Euro in Retail and E-Commerce: The Role of Institutional Framing, Trust, and Usability. *European Journal of Management, Economics and Business*, 3(3), 64-80.
4. Sipos, D. Conversational Artificial Intelligence in Retail: From Customer Support to Behavioral Influence.
5. Sipos, D. (2026). Conversational Artificial Intelligence in Retail: From Customer Support to Behavioral Influence. В *Artificial Intelligence for Resilient and Sustainable Technologies* (стр. 1-38). Deep Science Publishing.
6. Прокопенко, Ж. В. (2017). Господарська діяльність небанківських фінансових установ як об'єкт нормативно-правового регулювання: облікові аспекти. *Агросвіт*, (8), 60-64.
7. Панченко, И. А., & Прокопенко, Ж. В. (2013). Страховые резервы как объекты бухгалтерского учета страховых компаний: влияние условий риска. *Международный бухгалтерский учет*, (9), 53-63.
8. Прокопенко, Ж. (2013). Особливості облікового відображення формування технічних резервів у страховій компанії. *Бухгалтерський облік і аудит*, (3), 34-42.
9. Прокопенко, Ж. В., & Мельник, І. М. (2020). Механізм формування аналітичного забезпечення управління діяльністю лісгосподарських підприємств. *Економіка, управління та адміністрування*, (4 (94)), 69-73.
10. Прокопенко, Ж. В. (2012). Суб'єкти внутрішнього контролю в страхових компаніях. *Проблеми теорії та методології бухгалтерського обліку, контролю і аналізу*, (2 (23)), 290-296.