

Workers' wages in formal employment – AKM Model Estimates

Abstract

To supplement the information available to users of IMSS microdata, data from an AKM-type model estimation based on IMSS worker information are now available in the Econlab repository. The information contained in these databases is a valuable resource for complementing the economic analysis and research conducted at the DGIE. The datasets cover the period from 2004 to 2023 for both female and male workers. They include fixed effects associated with workers and firms.

Introduction

In recent years, the labor economics literature has focused on the AKM models proposed by Abowd et al. (1999). Using microdata from the Mexican Social Security Institute (IMSS), AKM models of the logarithm of wages were estimated. In the AKM model, the following linear specification is considered (Abowd et al. 1999; Card et al. 2003):

$$(1) \log(y_{it}) = \alpha_i + \psi_{J(i,t)} + X'_{it}\beta + r_{it}.$$

In Equation (1), y_{it} is the actual contribution base wage (in pesos as of July 2018) for worker i in month t . The fixed effects for individuals are represented by the parameters α_i , and those for the firm $J(i,t)$ where worker i is employed in month t are represented by $\psi_{J(i,t)}$. The vector X_{it} includes third-order polynomial terms for age and indicator variables by year. The error term in the model is denoted by r_{it} .

Individual fixed effects provide information on the component of wages attributable to worker characteristics; that is, they measure the contribution to individual wages of personal characteristics that do not vary over time, such as education once it has been completed. In turn, the component of wages attributable to firm characteristics is estimated through firm fixed effects, which capture the contribution to wages of firm

characteristics that do not vary over time, such as differences across industries or differences in productivity.

In order to correctly identify the fixed effects of workers and firms, the AKM model is estimated on the largest connected set—that is, on the sample of firms that are connected to one another through worker movements between them during the time window. After obtaining the fixed effects of workers who change firms, the fixed effects of workers who do not change jobs but belong to a firm in the largest connected set are imputed.

The model was estimated using three different samples. The first includes only male workers. The second includes only female workers. The third sample includes both male and female workers. Additionally, four time windows are considered: 2004–2008, 2009–2013, 2014–2018, and 2019–2023. Based on this, databases were constructed with fixed effects for workers and firms for each combination of sample and time window.

The model was estimated on three different samples. The first includes only male workers. The second includes only female workers. The third sample includes both male and female workers. It also considers four time periods: 2004–2008, 2009–2013, 2014–2018, and 2019–2023. Based on this, databases were constructed with fixed effects for workers and firms for each combination of sample and time period.

Variables

- **Employee ID:** Numeric variable containing unique identifiers for workers enrolled in the IMSS.
- **Employer ID:** Numeric variable containing unique identifiers for employer records enrolled in the IMSS.
- **worker_effects_2004_2008:** Numeric variable containing worker fixed effects from a model estimated for the 2004–2008 period.
- **worker_effects_2009_2013:** Numeric variable containing worker fixed effects from a model estimated for the 2009–2013 period.

- **worker_effects_2014_2018:** Numeric variable containing worker fixed effects from a model estimated for the 2014–2018 period.
- **worker_effects_2019_2023:** Numeric variable containing worker fixed effects from a model estimated for the 2019–2023 period.
- **workplace_effects_2004_2008:** Numeric variable with worker fixed effects from the model estimated for the 2004–2008 period.
- **workplace_effects_2009_2013:** Numeric variable with worker fixed effects from the model estimated for the 2009–2013 period.
- **workplace_effects_2014_2018:** Numeric variable with worker fixed effects from a model estimated for the 2014–2018 period.
- **workplace_effects_2019_2023:** Numeric variable with worker fixed effects from a model estimated for the 2019–2023 period.

Consultado el 5 de enero de 2026.
<https://www.inegi.org.mx/temas/uma/>

Keywords

AKM, matching, IMSS, labor market, wages.

DOI: STEF-AKM-EconLab

Additional Comments

The samples used to estimate Equation (1) are subject to the following restrictions:

- I) For each worker, only the job with the highest contribution base wage in each month is considered.
- II) Only workers aged 25 to 54 are included.

The variable of interest is the daily pre-tax wage..¹

It is important to note that contribution base wages greater than 25 Units of Measurement and Adjustment are censored..²

References

Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999). High Wage Workers and High Wage Firms. *Econometrica*, 67(2):251–333.

Card, D., Heining, J., and Kline, P. (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality. *The Quarterly Journal of Economics*, 128(3):967–1015.

Instituto Nacional de Estadística y Geografía (INEGI). 2025. Unidad de Medida y Actualización (UMA).

¹ The taxable daily wage includes various forms of compensation in addition to the base wage, such as paid vacation and bonuses. It does not include tax-exempt contributions.

² In 2018, the cap was 2,015.00 Mexican pesos (INEGI, 2025). We conducted a robustness analysis in which we estimated the AKM model using imputed wages for workers whose wages exceeded the wage cap. The results are similar.