

Recovering Latent Variables by Matching*

Manuel Arellano[†]

Stéphane Bonhomme[‡]

July 1, 2021

Abstract

We propose an optimal-transport-based matching method to nonparametrically estimate linear models with independent latent variables. The method consists in generating pseudo-observations from the latent variables, so that the Euclidean distance between the model's predictions and their matched counterparts in the data is minimized. We show that our nonparametric estimator is consistent, and we document that it performs well in simulated data. We apply this method to study the cyclicalities of permanent and transitory income shocks in the Panel Study of Income Dynamics. We find that the dispersion of income shocks is approximately acyclical, whereas the skewness of permanent shocks is procyclical. By comparison, we find that the dispersion and skewness of shocks to hourly wages vary little with the business cycle.

KEYWORDS: Latent variables, nonparametric estimation, matching, factor models, optimal transport, income dynamics.

JEL CODES: C14, C33.

*We thank Colin Mallows, Tincho Almuzara, Alfred Galichon, Jiaying Gu, Kei Hirano, Pierre Jacob, Roger Koenker, Thibaut Lamadon, Guillaume Pouliot, Azeem Shaikh, Tim Vogelsang, Daniel Wilhelm, and audiences at various places for comments. Tincho Almuzara and Beatriz Zamorra provided excellent research assistance. Arellano acknowledges research funding from the Ministerio de Economía y Competitividad, Grant ECO2016-79848-P. Bonhomme acknowledges support from the NSF, Grant SES-1658920.

[†]CEMFI, Madrid.

[‡]University of Chicago.