



Estimation with pairwise observations

E. Chan^a , **L. Mátyás**^b  and **K. Reizinger**^b

^a*School of Accounting, Economics and Finance, Curtin University, Perth, Western Australia*

^b*Department of Economics and Business, Central European University, Budapest, Hungary*
Email: F.Chan@curtin.edu.au

Abstract: The paper introduces a new estimation method for the standard linear regression model called *Estimation with Pairwise Observations* (EwPO). The procedure is not based on the optimisation of any objective function rather, it is a simple weighted average of slopes from observation pairs. Specifically, let (x_i, y_i) be the i^{th} observation drawn from the model

$$y_i = \beta_0 + \beta x_i + u_i, \quad i = 1, \dots, N \quad (1)$$

where $u_i \sim D(0, \sigma^2)$. The paper considers a set of slopes, $\{\beta_{ij}\}$ as defined by

$$\beta_{ij} = \frac{y_i - y_j}{x_i - x_j} = \frac{\Delta y_{ij}}{\Delta x_{ij}} \quad (2)$$

$\forall i, j = 1 \dots, N, i \neq j$ and investigates the relation between $\{\beta_{ij}\}$ and β . The paper focus on the estimator

$$\hat{\beta} = \left(\sum_{i=2}^N \sum_{j=1}^{i-1} w_{ij} \right)^{-1} \sum_{i=2}^N \sum_{j=1}^{i-1} w_{ij} \beta_{ij} \quad (3)$$

where w_{ij} are weights specified by the users. The paper derives the asymptotic distribution for $w_{ij} = \Delta x_{ij}$ and proposes a set of jackknife procedures to obtain the confident intervals for the case $w_{ij} = |\Delta x_{ij}|$. Monte Carlo simulations show that EwPO has good finite sample performance under the two weighting schemes.

Unlike traditional methods, such as Least Squares and Maximum Likelihood, among others, the estimated residual from EwPO is generally not by construction orthogonal to the explanatory variables of the model, therefore can be used for testing endogeneity, i.e., the correlation between the explanatory variables and the disturbance terms. The paper proposes a new test statistics for exogeneity based on EwPO and derives its asymptotic distribution. The distribution is a Brownian functional and therefore the critical values must be simulated.

In addition, for certain specification, there exists a debiased procedure for EwPO which enables consistent estimation of the unknown slope parameters of the model in the presence of endogeneity without relying on additional external information usually needed in the form of instrumental variables or moment conditions.

Monte Carlo simulations show that the test statistics has good power against different alternatives, different levels of correlation.

Keywords: *Linear regression model, consistent estimation, endogeneity, testing for endogeneity, instrumental variables*