Table 3: Summary statistics and variable descriptions

Variable	Mean	Standard deviation	N	Description
Per capita CO_2	3.82	4.73	232	Carbon dioxide emissions, industrial, in metric tons per capita
Deforestation	0.68	1.28	223	Annual deforestation, average percentage change, 1990-1995
Energy depletion	3.13	7.43	223	In percent of GDP, equal to the product of unit resource rents and the physical quantities of fossil fuel energy extracted
Rural water access	51.20	27.42	137	Access to clean water, percentage of rural population, 1990-1996
Urban water access	76.28	21.76	140	Access to clean water, percentage of urban population, 1990-1996
GATT/WTO membership	0.78	0.41	232	Member country of GATT/WTO: 1 if member, 0 otherwise
Real GDP per capita	7.30	7.47	232	Real (1990) gross domestic product per capita, in thousands of dollars
Polity	3.17	6.85	232	Index, ranging from-10 (strongly autocratic) to 10 (strongly democratic)
Area per capita	51.6	89.56	232	Land area divided by population

Note: Same data used by Millimet and Tchernis (2009). We thank Prof. Millimet for sharing the data with us. Original source is Environmental indicators and country-level controls are from Frankel and Rose (2005), whereas GATT/WTO membership data are from Rose (2004). N = number of observations. For further details, see the aforementioned papers.

for each outcome subsample. More specifically, for each outcome, we model the probability of a country being a GATT/WTO member ($D = 1$ if member, $D = 0$ otherwise) by a standard Probit model and consider two different specifications:

Spec1: X includes real per capita GDP, land area per capita, and polity.

Spec2: X is defined as in *Spec1* but adds pairwise interaction terms between each covariate.

For each of these specifications, we test the null hypothesis

$$H_0 : \exists \theta_0 \in \Theta : \mathbb{E}\left[D - \Phi\left(X'\theta_0\right) | \Phi\left(X'\theta_0\right)\right] = 0 \ a.s.,$$

against H_1, which is simply the negation of H_0. Table 4 reports the testing results for each specification, together with normalized IPW estimator for the ATE based on (4.3). We also consider the normalized IPW estimator for the ATT,

$$\widehat{ATT}_n = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{w_{1,i}^{treat}}{\bar{w}_{1,n}^{treat}} - \frac{w_{0,i}^{treat}}{\bar{w}_{0,n}^{treat}}\right) Y_i, \tag{5.1}$$

where $w_{1,i}^{treat} = D_i$, $w_{0,i}^{treat} = (1 - D_i)\, q\left(X_i, \hat{\theta}_n\right) / \left(1 - q\left(X_i, \hat{\theta}_n\right)\right)$, and $\bar{w}_{d,n}^{treat}$ is the sample mean of $w_{d,i}^{treat}$, $d = \{0, 1\}$. The associated standard errors and p-values are in parenthesis and brackets, respectively. Following Millimet and Tchernis (2009), we trim observations with estimated propensity score outside the interval $[0.05, 0.95]$ to avoid denominators arbitrarily close to zero. Bootstrapped p-values for our proposed specification tests are based on 100,000 bootstrap draws[7].

At the 5% level we find that, based on the CvM_n test statistic (2.8), *Spec1* is rejected for per capita CO_2, deforestation and energy depletion, but is not rejected for rural and urban

7 Note that the variables in *Spec1* and *Spec2* are all functions of the same three covariates: real per capita GDP, land area per capita, and polity. As so, *Spec1 and Spec2* have the same information content with respect to the reliability of Assumption 2.1.