

or World Trade Organization (WTO) on 5 different measures of environmental quality: per capita  $CO_2$  emissions, average annual deforestation rate for 1990-1996, energy depletion, rural access to clear water and urban access to clear water.

As [Millimet and Tchernis \(2009\)](#), we use three different covariates to model the probability of being a GATT/WTO member: real per capita GDP, land area per capita, and polity, which is a measure of how democratic (versus autocratic) is the structure of the government. The motivation to include these three covariates is to increase the plausibility of Assumption 2.1 as discussed by [Frankel and Rose \(2005\)](#). GDP per capita is associated with the probability of being member of the GATT/WTO and at the same time may have effects on different measures of environment quality, for instance, via the environmental Kuznetz curve. Land area per capita is another potential confounder since higher population density may lead to environmental degradation and “larger” countries are more likely to trade more, affecting the probability of being member of the GATT/WTO. Finally, as noted by [Frankel and Rose \(2005\)](#), low-democracy countries tend to have lower measures of environmental quality, and can also confound the effect of GATT/WTO membership. In what follows and in the same spirit of [Millimet and Tchernis \(2009\)](#), we assume that Assumption 2.1 holds after controlling for these three confounding factors<sup>6</sup>.

The unbalanced country-level panel data we use follows from [Millimet and Tchernis \(2009\)](#), and includes observations from 1990 (before the WTO) and 1995 (after the WTO). However, it is important to have in mind that treatment is defined as being a GATT/WTO member, and therefore there are countries who were treated in both times, and others who were treated only in 1995. Table 3 provides summary statistics and more detailed description of the variables. Finally, we highlight that the data we analyze is from an unbalanced country-level panel and instead of only considering the “always observed” countries, we follow [Millimet and Tchernis \(2009\)](#) and run a separate analysis for each outcome. For further details, see Section 4.1 of [Millimet and Tchernis \(2009\)](#).

The main goal of this section is to assess the “reliability” of different treatment effect measures by analyzing if different propensity score models are correctly specified or not. Given that the sample constitutes of an unbalanced panel, we estimate separate propensity score models

---

6 If one finds the plausibility of this assumption rather low, all the estimates presented below should be interpreted as associations/correlations and not as causal effects. In light of Remark 1, however, Assumption 2.1 plays no role in our specification tests.