

DATASET: ForeignPresence&ETCs.dta

VARIABLE LABELS AND DEFINITIONS

PROVINCE-LEVEL VARIABLES:

- MNC (MNC Activity): the first principal component of inward FDI (% of GDP) and total trade of foreign invested enterprises (% of GDP) in 2004.
- lwdist (Geographic Closeness): weighted geographic distance (log).
- lgdpper2003: GDP per capita in 2003 (log).
- gdp2003: total GDP in 2003.
- lpop2003: total population in 2003 (log).
- province: province name.
- provinceid: province ID.
- dup_prov_first: first observation of province-level duplicates (dummy).

TWO-DIGIT INDUSTRY-LEVEL VARIABLES:

- foreign_output (Foreign Presence): output share of foreign firms.
- HMT_output (Foreign Presence, HMT Origin): output share of foreign firms originated in Hong Kong, Macao, and Taiwan.
- NonHMT_output (Foreign Presence, Non-HMT Origin): output share of foreign firms originated in countries other than Hong Kong, Macao, and Taiwan.
- w_con_pdc4 (CR4): weighted four-firm concentration ratio (output share of the largest four firms in the industry).
- w_con_pdc8 (CR8): weighted eight-firm concentration ratio (output share of the largest eight firms in the industry).
- marketsize (Market Size): log of total value added.
- growthrate (Growth Rate): growth rate of total output from 2001 to 2003.
- last_intensity (Asset Intensity): log of fixed assets per employee.
- lscale (Scale Economies): log of the average size of the largest firms that account for 50% of total output in the industry.
- mgovthelp (Government Help): industry median of a linear combination of the percentages of the officials in government departments (taxation, public security, environment, and labor and social) who are perceived to contribute to the development of the company.
- mtaxrate (Tax Burden): industry median of firms' tax rates (total taxes divided by total sales).
- prov_ind: province-industry identification code.
- dup_first: first observation of province-industry duplicates (log).

FIRM-LEVEL VARIABLES

- idstd: unique firm id.
- letcs (ETCs): firm expenditure on entertainment and travel costs as % of total revenue (log).
- soe (SOEs): a dummy variable coded 1 if firms are registered as state-owned and the share of state capital is equal to or greater than 50%, and 0 if otherwise.

- collective (Collective): a dummy variable coded 1 if firms are registered as collective-owned and the share of collective capital is equal to or greater than 50%, and 0 if otherwise.
- private (Private): a dummy variable coded 1 if firms are registered as private-owned and the share of private capital is equal to or greater than 50%, and 0 if otherwise.
- foreign_HMT (Foreign, HMT): a dummy variable coded 1 if firms are registered as foreign-owned (Hong Kong, Macao, and Taiwan) and the share of foreign capital is equal to or greater than 50%, and 0 if otherwise.
- foreign_NonHMT (Foreign, Non-HMT): a dummy variable coded 1 if firms are registered as foreign-owned (other than Hong Kong, Macao, and Taiwan) and the share of foreign capital is equal to or greater than 50%, and 0 if otherwise.
- mixedown (Mixed ownership): a dummy variable coded 1 if firms are not coded as either SOEs, collective, private, or foreign, and 0 if otherwise.
- revenue (Revenue): log of total business income (core business income + other business income) in 2004.
- lemp (Employees): log of total employees in 2004.
- sales_otherprov (Out-of-Province Sales): a dummy variable coded 1 if firms sell to other provinces, and 0 if otherwise.
- govtsales (Proportion of Sales to Government): Proportion of sales to government in 2004.
- soesales (Sales to SOEs): Proportion of sales to SOEs in 2004.
- relationship (Years of Relationship): total years of firms' relationship with their major clients and suppliers.
- licenses (Licenses): number of licenses and registrations (permanent and renewable annually) required for firms.
- lceopay (CEO Pay): log of CEO annual income. This variable is not directly observed. The survey reports workers' average wage, as well as the ratio of CEO's annual income to the mid-level managers', and the ratio of the mid-level managers' annual income to the ordinary employees'. CEO annual income is calculated using the product of workers' average wage and the two ratios.
- interaction (Interaction with government): number of days that a firm's general manager (GM) or vice GM spends on government assignments and communications.
- gm_govt (Government Appointed General Manager): a dummy variable coded 1 if the firm's general manager is appointed by the government, and 0 if otherwise.