Sampling methodology

The survey has been carried out in the following 9 provinces: Hanoi, Hai Phong, Ho Chi Minh City, Phu Tho, Nghe An, Quang Nam, Khanh Hoa, Lam Dong and Long An.

The population of non-state manufacturing enterprises in the selected provinces is based on two data sources from the General Statistics Office of Vietnam (GSO): The Establishment Census from 2002 and the Industrial Survey 2004-2006. The number of individual business establishments that do not satisfy the conditions stated in the Law on Enterprises of Vietnam (also referred to as household enterprises) are obtained from the Establishment Census. This information is combined with data on enterprises formally registered under the Law on Enterprises at the province level from the Industrial Survey (GSO, 2003). This provides additional information on private, collectives, partnerships, private limited enterprises, and joint stock enterprises. Joint ventures have been excluded from the sampling framework due to the often unclear nature of government and foreign involvement in such ownership structures. Stratified sampling was used to ensure adequate number of enterprises in each province with different ownership forms. It is important to note that the sample of informal firms is not representative of the informal sector as a whole in Vietnam, since the sampling scheme of the survey is based on the GSO business censuses and surveys, which cover only part of the informal sector.

The enterprises are distributed across approximately eighteen sectors and the dominant ones are: food processing, fabricated metal products, and manufacturing of wood products. Enterprises are classified as micro, small, medium, and large according to the current World Bank definition. Micro-enterprises have up to 10 employees, small-scale enterprises up to 50 employees, medium-sized enterprises up to 300 employees, while large enterprises have over 300 employees.

The employee module collects data from 1342 employees from 524 randomly selected enterprises. Note that the employee data are not representative.