Description of Data and Code Files for
"Do Multinationals Transfer Culture? Evidence from Female Employment in China"
Journal of International Economics

Heiwai Tang and Yifan Zhang (2021)

A. Instructions on how to obtain the proprietary data

The Annual Survey of Industrial Firms, China Customs Transaction-Level Data, the 2005 Mini Population Census, Firm Patent Database and Foreign Invested Firms Survey are *proprietary* data maintained by the Chinese government. We are not allowed to disseminate these data.

However, the raw data on industrial firm survey, the customs data, and the population census can be purchased from:

> Beijing Fu'ao Huamei Information Consulting Co., Ltd.
> Website: http://www.allmyinfo.com/
> Tel: 8610-63754526. Fax: 8610-63751229.

The firm patent database can be purchased from:

> Beijing INCOPAT CO., LTD
> Website: www.incopat.com
> Tel: 8610-60607720-619. Fax: 8610-60607720-647

B. Description of the Stata do files

1. *data construction.do*

- Assemble the datasets from NBS, customs, census, FIE survey and firm patents.
- Generate the main dataset "main_panel.dta".

2. *all regressions.do*

- Generate Figure 1 and Figure 2.
- Run regressions of cultural transfer, profitability and cultural spillovers.
- Generate regression results reported in Table 2, Table 3, Table 4 and Table 5.

3. *misallocation.do*

- Compute the extent of misallocation due to gender discrimination.
- Compute the TFP gains arising from increased female employment induced by FDI cultural spillover.
- Generate results reported in Figure 3, Table 6 and Table 7.

C. Description of the Stata data files

1. *country_level_gii.dta*

   - data for country level gender inequality index in 2011.

2. *country_level_gdppc.dta*

   - data for country level GDP per capita in 2004.

3. **FIE_survey_country_code.dta**

   - data for the concordance table of the country code in the FIE surveys.

4. *herf.dta*

   - data of the Herfindahl index at the four-digit industry level.

5. **HS02-CIC03.dta**

   - concordance table between HS 2002 product code and CIC 2003 industry code.

6. **HS07-HS02.dta**

   - concordance table between HS 2002 product code and HS 2007 product code.

7. *import_output_ratio_cic.dta*

   - import value to output ratio by four-digit industry.

D. Other files

In the paper, when we calculated real capital stock and productivity of industrial firms, we used some files that were constructed by the following paper:

Brandt, L., J. Van Biesebroeck, and Y. Zhang (2012). "Creative Accounting or Creative Destruction? Firm-level Productivity Growth in Chinese Manufacturing," *Journal of Development Economics*, 97(2):339-351.

The information is available online at
http://feb.kuleuven.be/public/N07057/CHINA/appendix/