Total factor productivity in South African manufacturing firms

Supporting information

Appendix A Database construction

Compiling the CIT database from the raw data involves a number of steps. During the sample period in question, the format of CIT returns changed. Specifically, SARS changed the submission form from what was called the IT14 form to the ITR14 form. This change came into effect on 4 May 2013 (SARS 2016). The main difference between the IT14 and the ITR14 is the depth of data submission required of companies of different sizes. All firms, regardless of size, were required to submit the full IT14 form whereas for ITR14 the level of detail that firms are required to submit varies depending on firm size. This change in the way data are gathered affects the way in which the variables used for our analysis are constructed. This is discussed in section 2..

The IRP5 is a reconciliation form that includes details of the total amount paid to employees from different sources, the total amount of employer's tax that is paid, skills development levy payments, unemployment insurance fund payments, employment tax incentives deducted, as well as the periods worked in the year of assessment. An employer must issue an IRP5 certificate to each employee to whom remuneration is paid or has become payable and from which employees' tax for a given tax period has been deducted.[1] Where no employee tax has been deducted from remuneration, an IT3(a) form is submitted to an employee. Where an employee earns less than R2,000 in a given tax year and where no employee tax has been deducted from remuneration paid, the employee is not given an IRP5 or an IT3(a) form. The tax year runs from 1 March in a previous calendar to the last day of February in the tax year.

The IT14 and ITR14 forms are submitted by firms for a tax year determined by their financial year-end. That is, a firm with financial year-end on 31 January 2012 will submit financial statements for the financial year 1 February 2011 to 31 January 2012 for the tax year ending on 29 February 2012. Similarly, where a firm has a financial year from 1 January 2012 to 31 December 2012, the firm will be assessed for the tax year ending on 29 February 2012. IRP5 data, on the other hand, report on labour in the firm for the year from 1 March to 28 February the following year, regardless of the financial year-end. Although for most firms the financial year coincides with the SARS tax year, around 15 per cent of firms do not follow SARS conventions. To ensure that firms are reporting as close as possible in the same financial year, we move firms with financial year ending after 30 August to the next financial year. Thus, a firm with financial year ending on 31 September 2012 will be moved from tax year 2012 to tax year 2013. This also allows us to make sure that the firm's labour data capture most of the activity occurring during a specific year.[2]

Fixed capital is reported differently in IT14 and ITR14. In IT14, all firms are asked to report their fixed property, their fixed assets, and their other fixed assets. In ITR14, firms of different types are asked different questions. Micro firms and small firms (as defined by their entry type in the online questionnaire) report fixed assets in a single item that combines property, plant, and equipment. Medium to large firms, on the other hand, are asked to submit three line items: fixed property, fixed assets, and other fixed assets. For ITR14, we construct fixed assets sum of these three line items where the missing entries are treated as zeros.

In the IT14 form, sales are reported as turnover in a single line item. In the ITR14 form, different firms submit different measures. Micro and small firms report sales or turnover in the same way

[1] Paragraph 13(1) of Schedule 4 of the Income Tax Act (Government of South Africa 1962).

[2] Note that the IRP5 data actually contains information on the exact day a person is employed to the date employment ceases; however, data for 2009 and 2010 were not available.

as in the IT14 form, whereas medium to large firms submit foreign connected sales and other sales.

Appendix B Productivity estimation

We assume a Cobb–Douglas production function written in the following form for the purpose of empirical estimation:

$$y_{it} = \beta_l l_{it} + \beta_k k_{it} + \omega_{it} + e_{it}, \quad \text{(A1)}$$

where y_{it} is the log of value added, l_{it} is the log of the labour input, k_{it} is the log of the capital input, ω_{it} is unobserved productivity, and e_{it} is an unanticipated shock or random error term.

The theory underlying the Levinsohn and Petrin (2003) approach to estimating productivity is that for some function $f(.)$

$$\omega_{it} = f\left(k_{it}, m_{it}\right), \quad \text{(A2)}$$

where m_{it} are intermediate inputs.[3]

As in the Olley and Pakes (1996) approach, we assume that productivity evolves according to a first-order Markov process. So:

$$E\left(\omega_{it} \middle| \omega_{it-1}, \omega_{it-2}, ..., \omega_{i1}\right) = E\left(\omega_{it} \middle| I_{it-1}\right) = E\left(\omega_{it} \middle| \omega_{it-1}\right), \quad \text{(A3)}$$

where I_{it-1} is the information set at time $t-1$ and all past realizations of productivity are assumed to be part of that information set. In other words, the firm expectations about future productivity depend only on the productivity in the previous period.

We assume that labour is chosen at the same time that productivity is realized but that intermediate inputs and capital stock k_{it} are determined at time $t-1$.

Assuming that $E(e_{it} | k_{it}, m_{it})=0$, and substituting for ω_{it}, the production function in Equation (A1) can be written as:

$$y_{it} = \beta_l l_{it} + \beta_k k_{it} + f\left(k_{it}, m_{it}\right) + e_{it}, \quad t = 1, 2, ..., T. \quad \text{(A4)}$$

The parameters β_l and β_k will not be separately identified, the former owing to collinearity between labour and productivity (Ackerberg et al. 2006) and the latter owing to the inclusion of k_{it} in $f(.)$.

Returning to the process assumed to underlie the evolution of productivity described in Equation (A3), we define innovation as follows:

$$x_{it} = \omega_{it} - E\left(\omega_{it} \middle| \omega_{it-1}\right). \quad \text{(A5)}$$

Combined with Equation (A2), which implies that $\omega_{it-1}=g(k_{t-1}, m_{it-1})$, and after some rearranging, Equation (A5) can be rewritten as:

---

[3] In the Olley and Pakes (1996) approach, investment is used in place of intermediate inputs.

$$w_{it} = f\left(k_{i=1} m_{it-1}\right) + x_{it}. \quad \text{(A6)}$$

Substituting into Equation (A1) provides us with a second equation that can be used to identify the two parameters of interest, β_l and β_k:

$$y_{it} = b_l l_{it} + b_k k_{it} + f\left(k_{it-1}, m_{it-1}\right) + v_{it}, \quad t = 2,3,...,T, \quad \text{(A7)}$$

where $v_{it}=\xi_{it}+e_{it}$. A set of suitable moment restrictions emerges from the assumptions underlying the evolution of productivity and the timing of the choice of inputs. Equation (A5) implies that innovation will be independent of the information set at time $t-1$ (i.e. ω_{t-1}). As k_{it} is determined at period $t-1$, it will be uncorrelated with unobserved innovation ξ_{it}. In other words:

$$E\left(x_{it} \middle| k_{it}\right) = 0. \quad \text{(A8)}$$

Innovation will, however, be correlated with any production decisions that are made between period $t-1$ and t. As such, the labour input, determined at period t, will be correlated with ξ_{it}. The lag of labour, l_{it-1}, however, will not, given that it is part of the information set at time $t-1$. As such:

$$E\left(x_{it} \middle| l_{it-1}\right) = 0. \quad \text{(A9)}$$

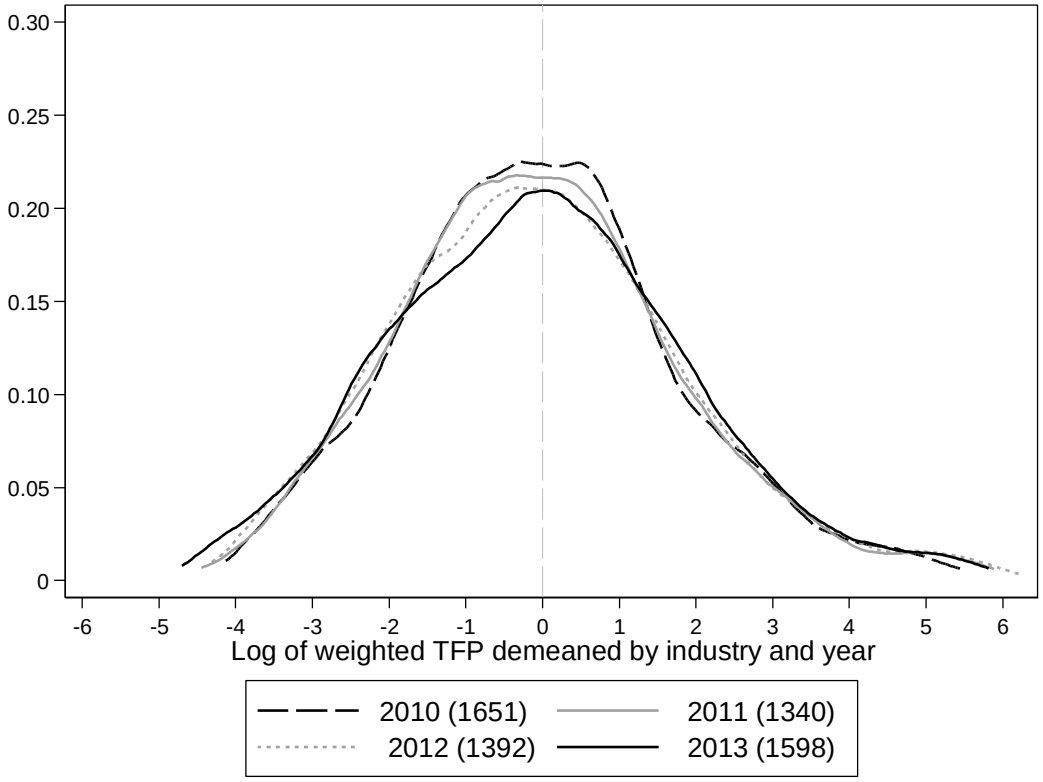
The full set of moment conditions is therefore given by $E(v_{it} \mid k_{it}, l_{it-1}, k_{it-1}, m_{it-1})=0$. The unknown functions $f(.)$ and $g(.)$ are approximated by third-degree polynomials. Equation (A7) can be estimated using pooled instrumental variables estimation with the instrument set $z_{it}=(k_{it}, l_{it-1}, k_{it-1}, m_{it-1}, \ldots)$, where all higher-order terms and their interactions in the polynomials act as their own instruments and all lags can also be used as instruments in testing overidentifying restrictions. In the estimation of Equation (A7), a full set of time dummies is included to control for heterogeneity over time in the production function and productivity. Once we have consistent estimators for β_l and β_k, productivity can be estimated using Equation (A10):

$$\hat{w}_{it} = y_{it} - \hat{b}_l l_{it} - \hat{b}_k k_{it}. \quad \text{(A10)}$$

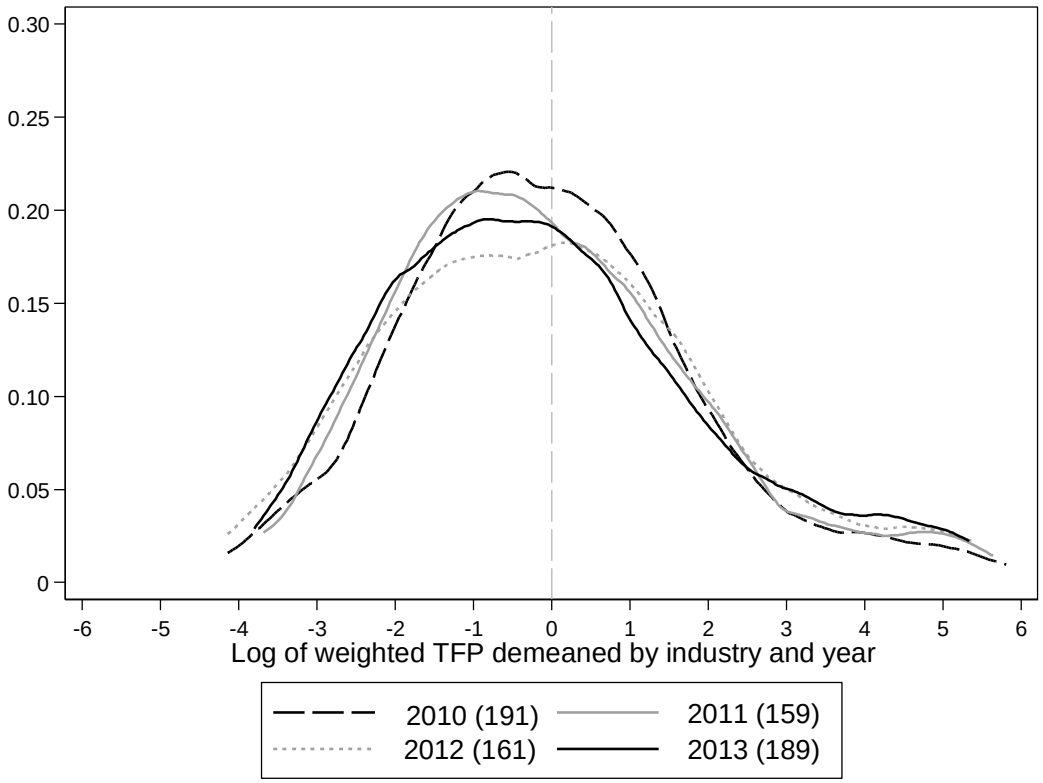
Appendix C Distribution of total factor productivity (TFP)

Figure C1: TFP distribution per industry per year

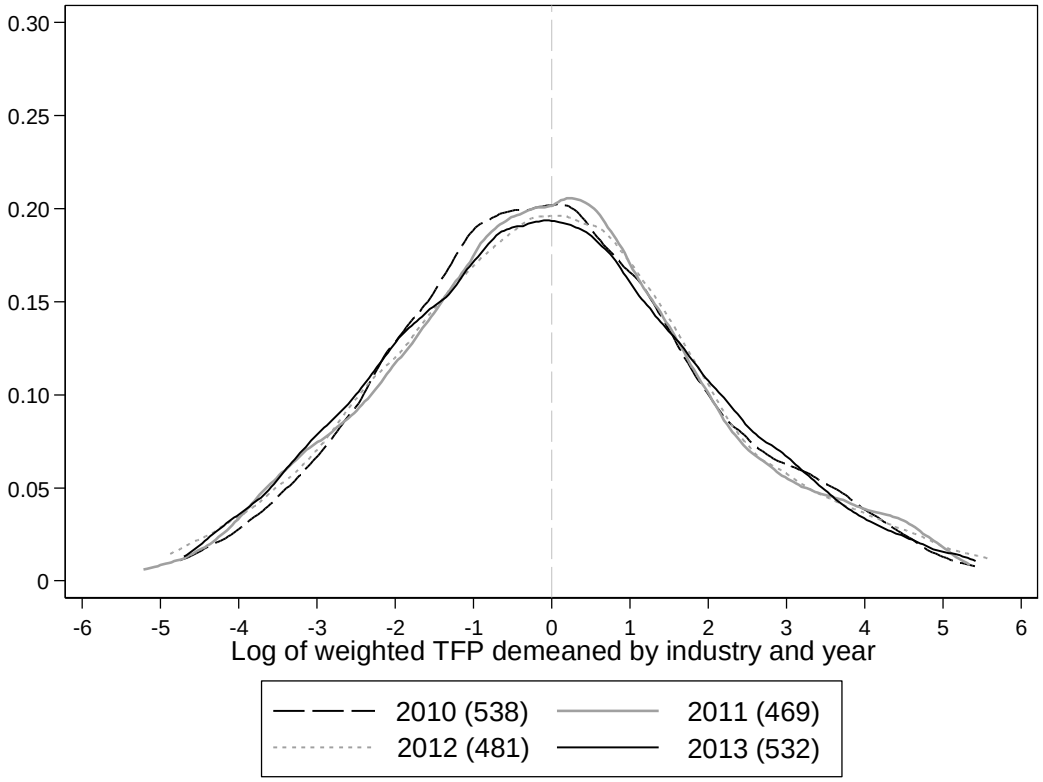
(a) 10: Food products



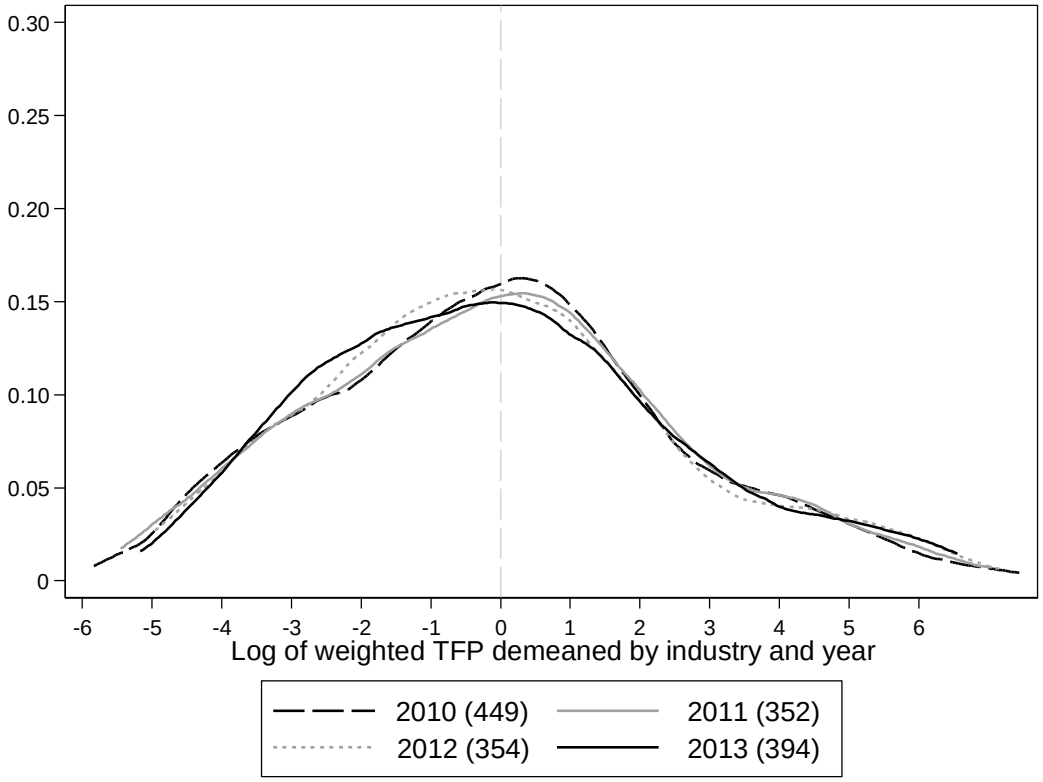
(b) 11: Beverages



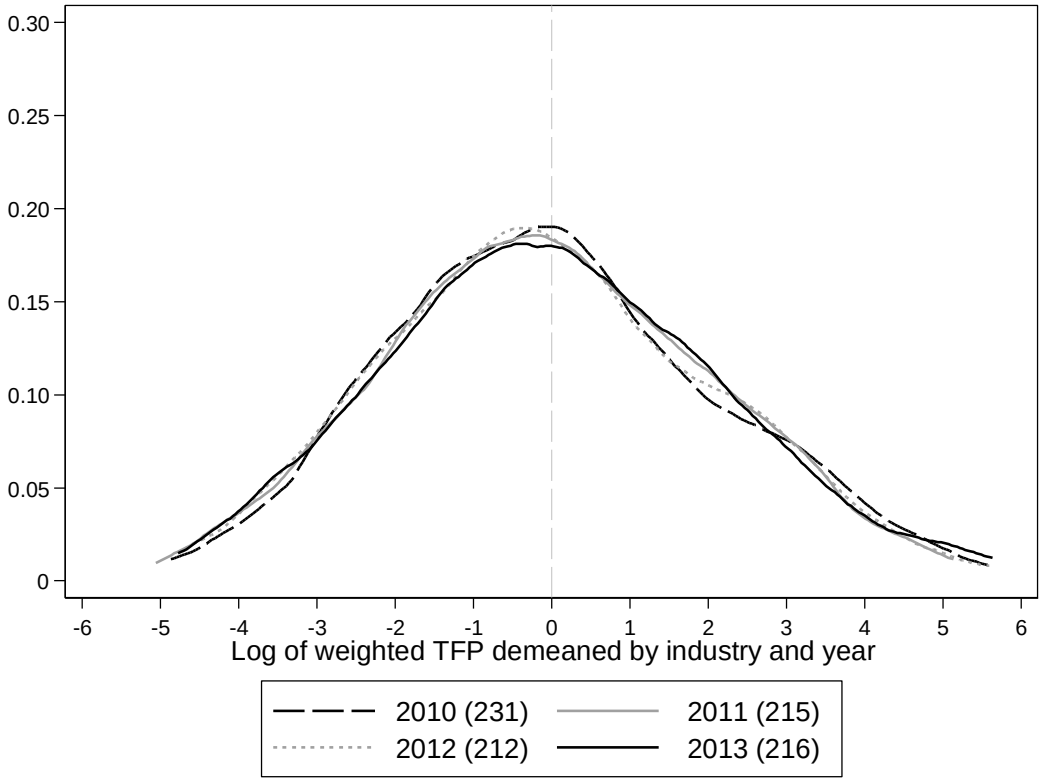
(c) 13: Textiles



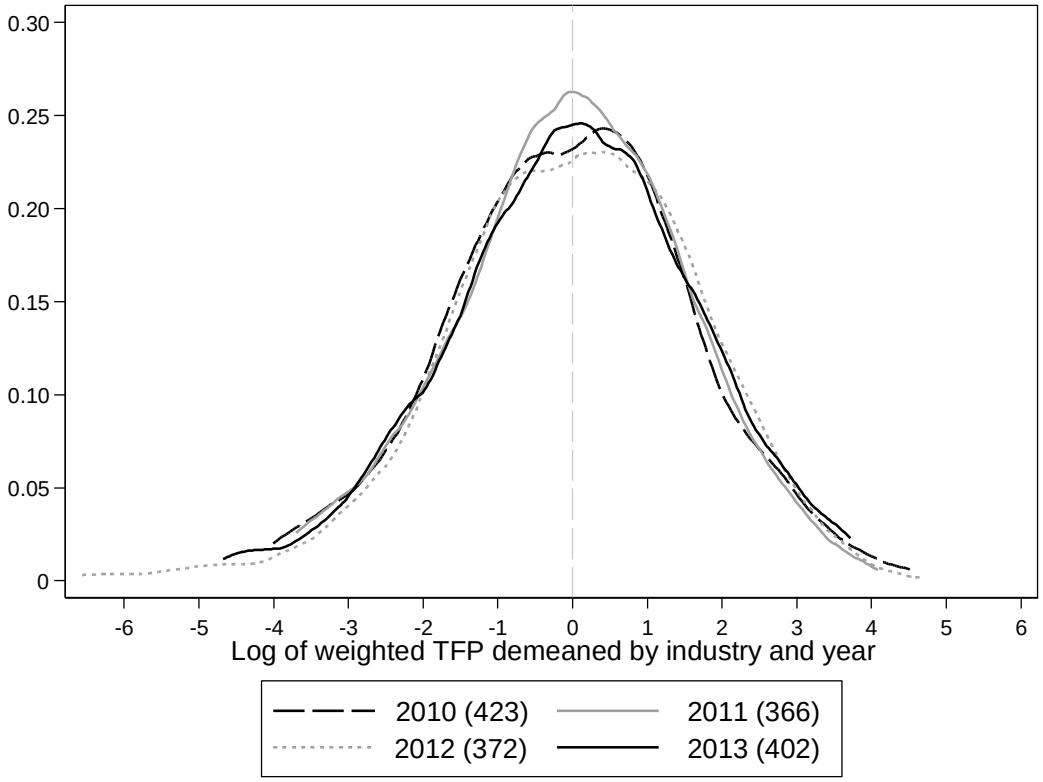
(d) 14: Wearing apparel



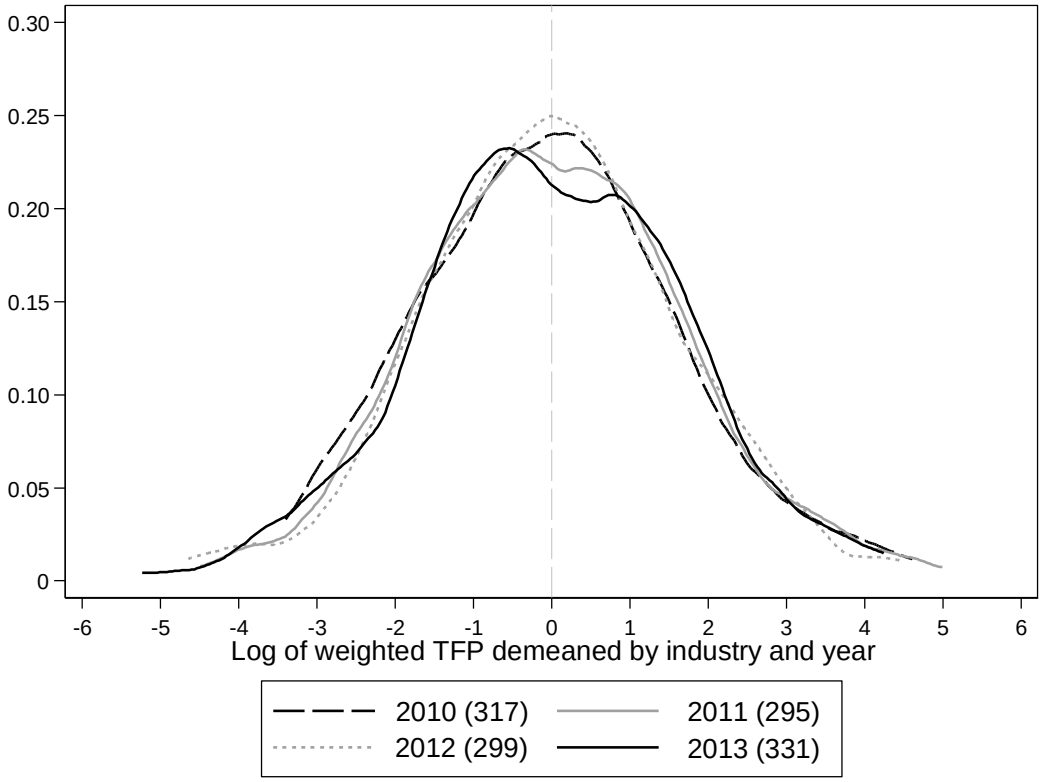
(e) 15: Leather and related products



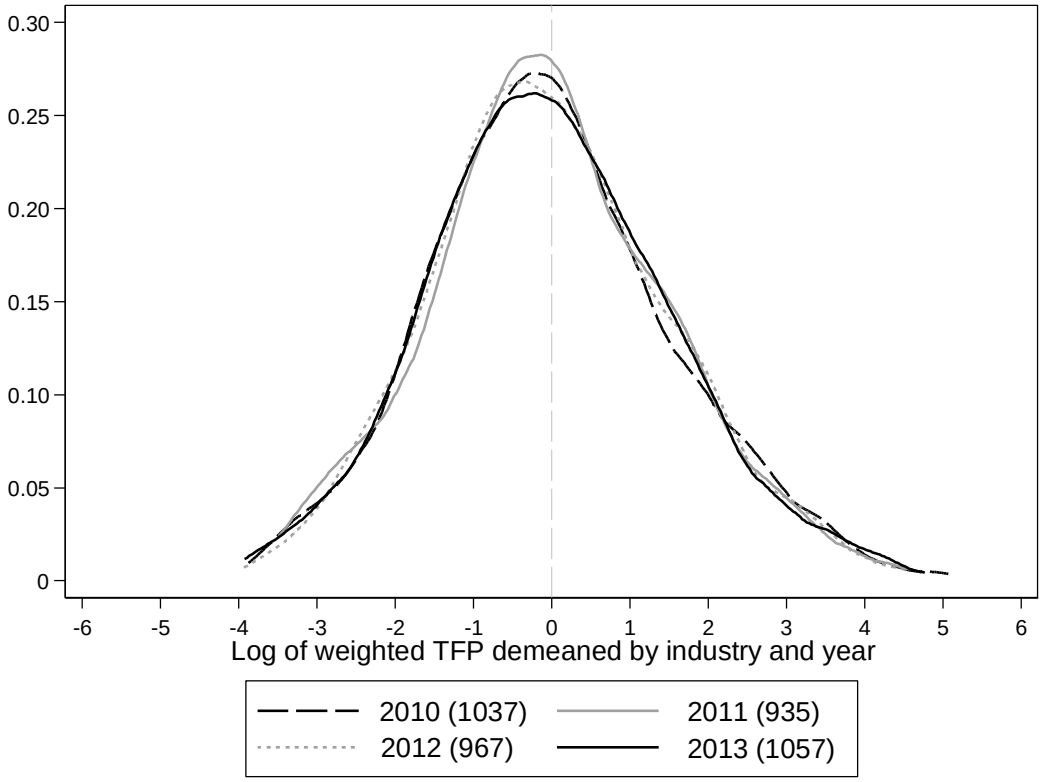
(f) 16: Wood and products of wood and cork, except furniture



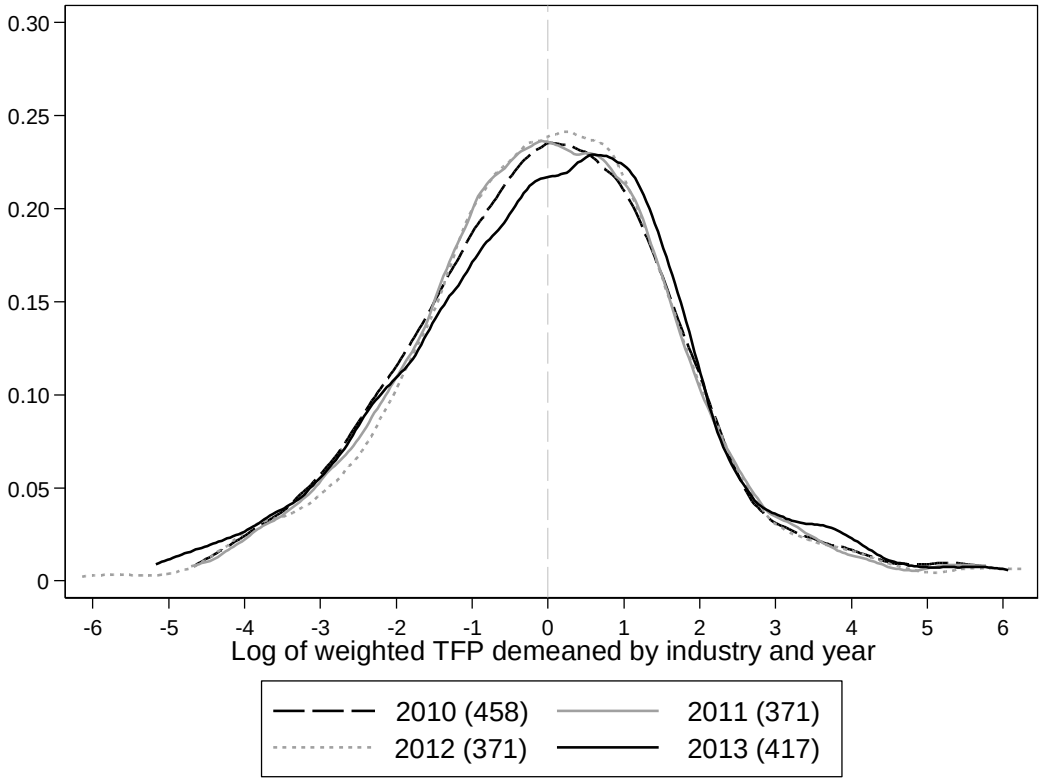
(g) 17: Paper and paper products



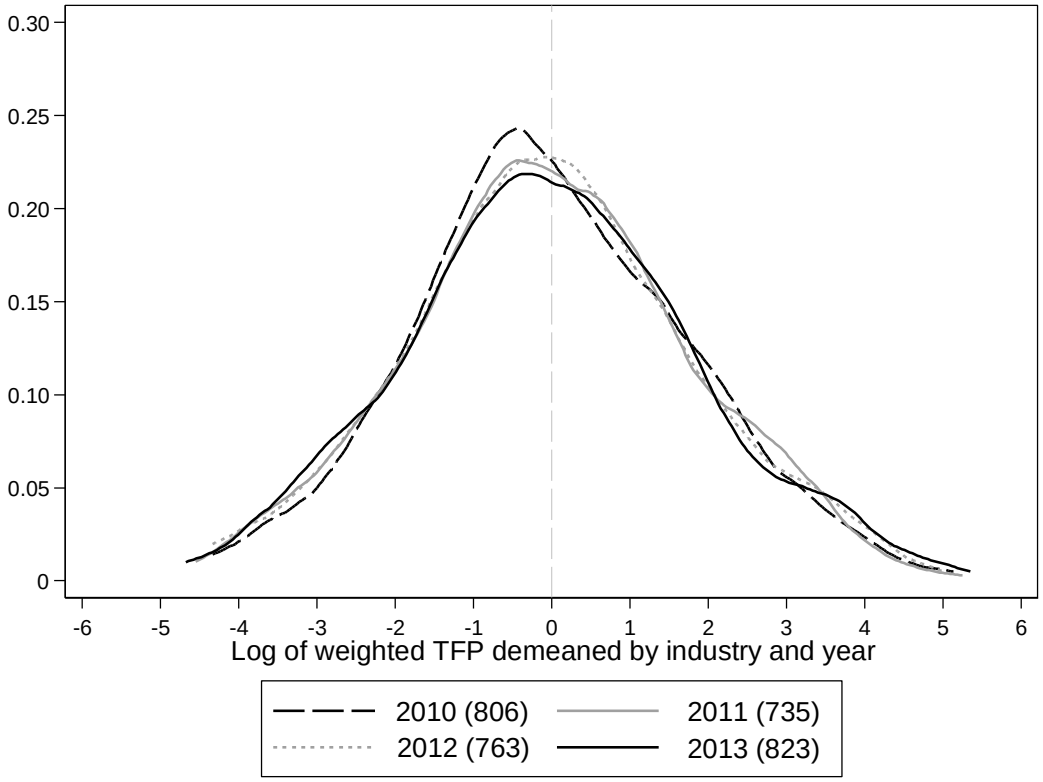
(h) 18: Printing and reproduction of recorded media



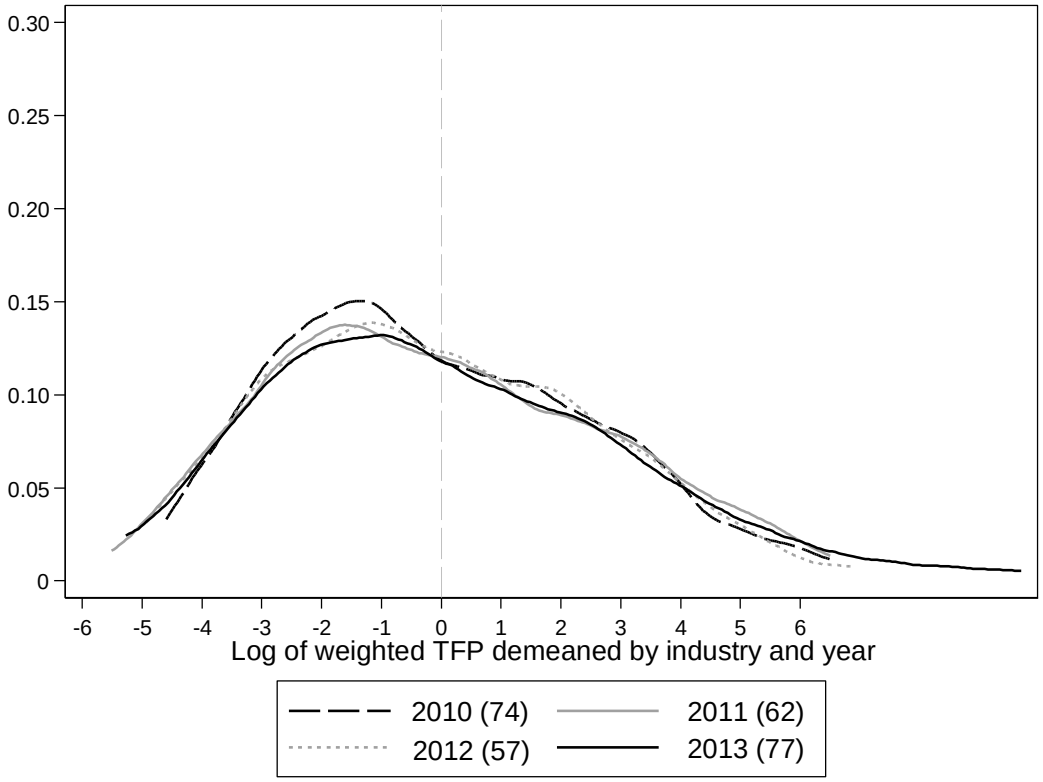
(i) 19: Coke and refined petroleum products



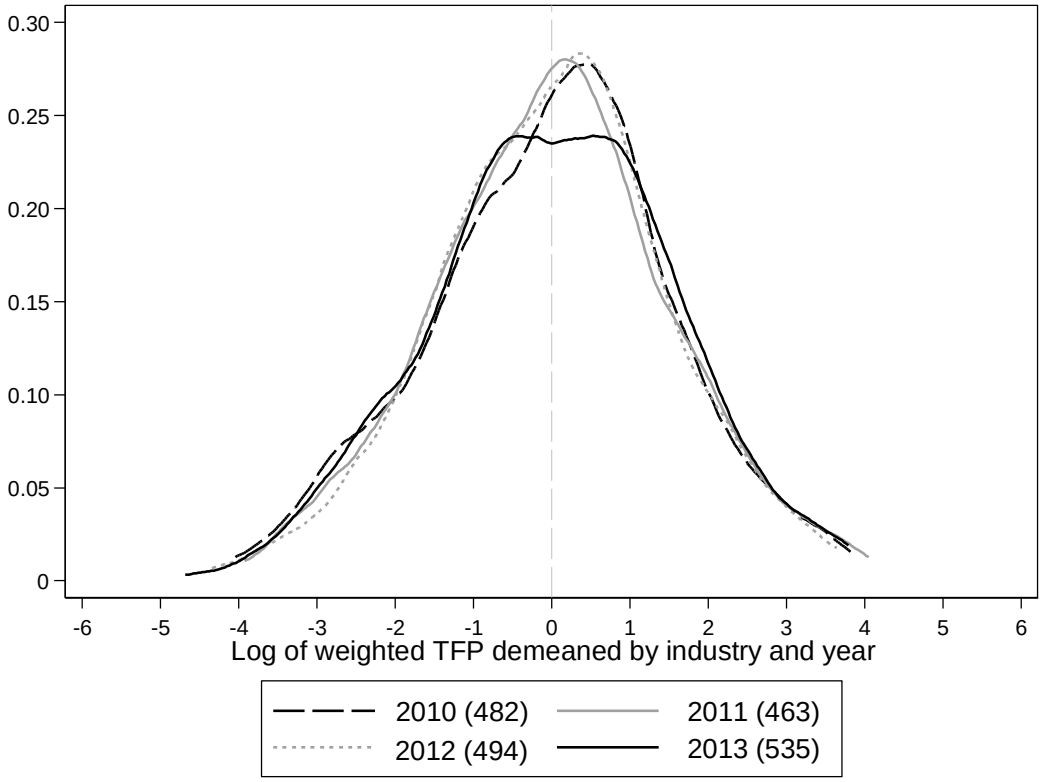
(j) 20: Chemicals and chemical products



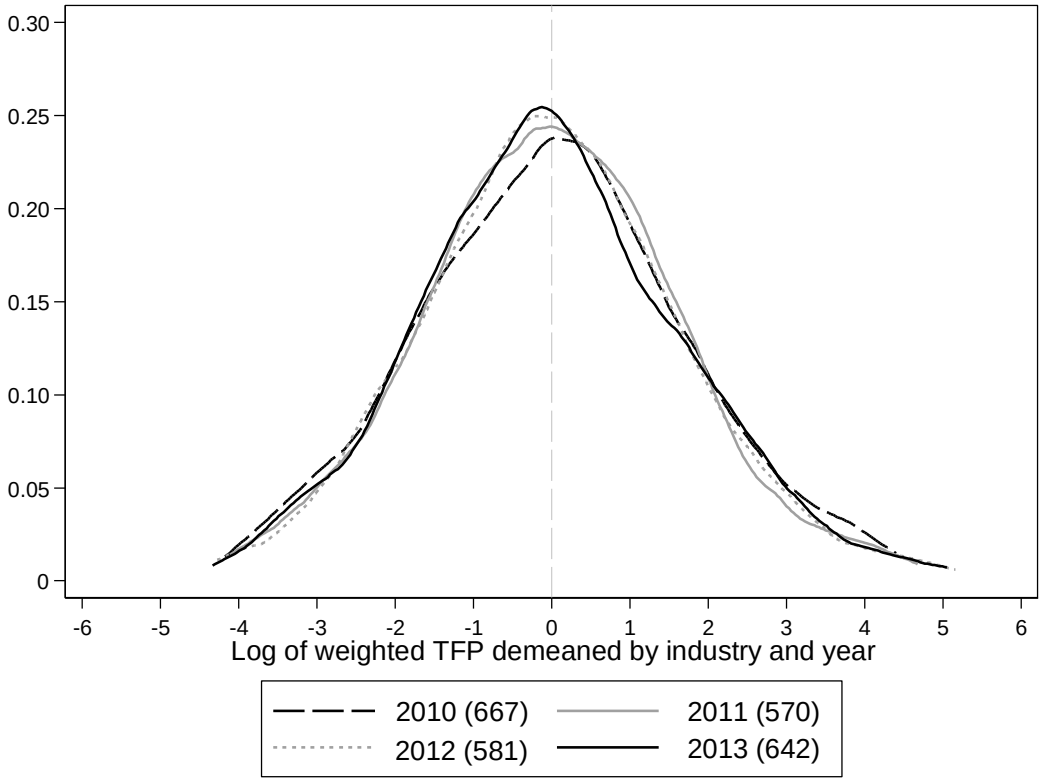
(k) 21: Basic pharmaceutical products and pharmaceutical preparations



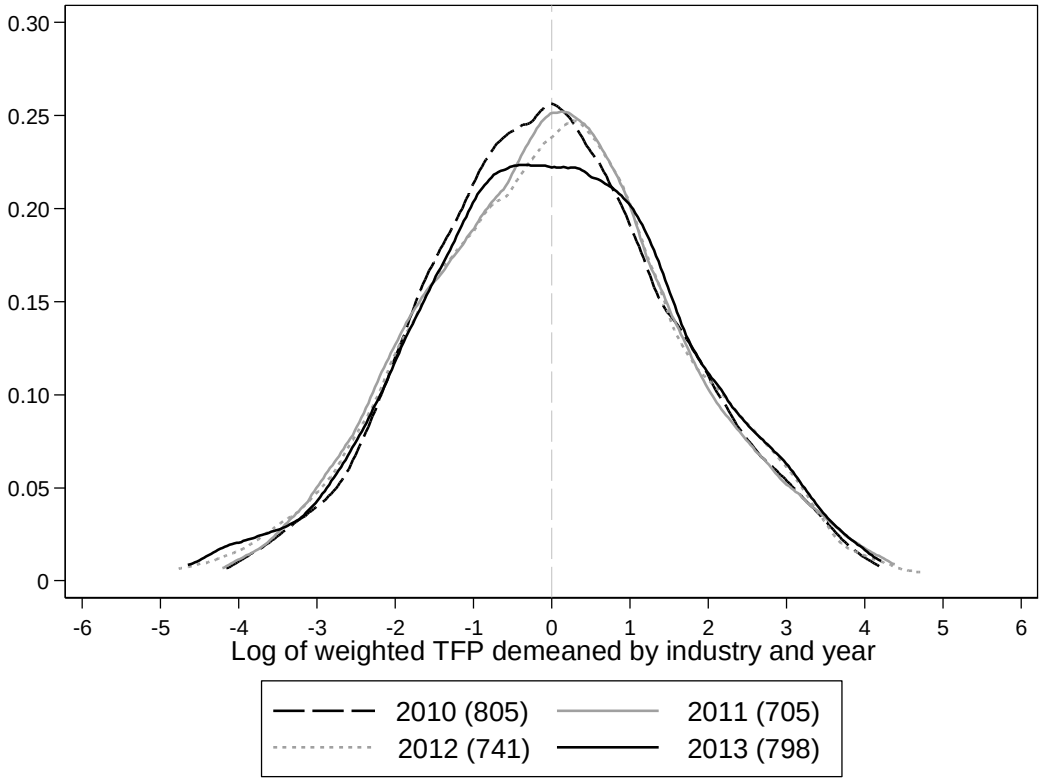
(l) 22: Rubber and plastics



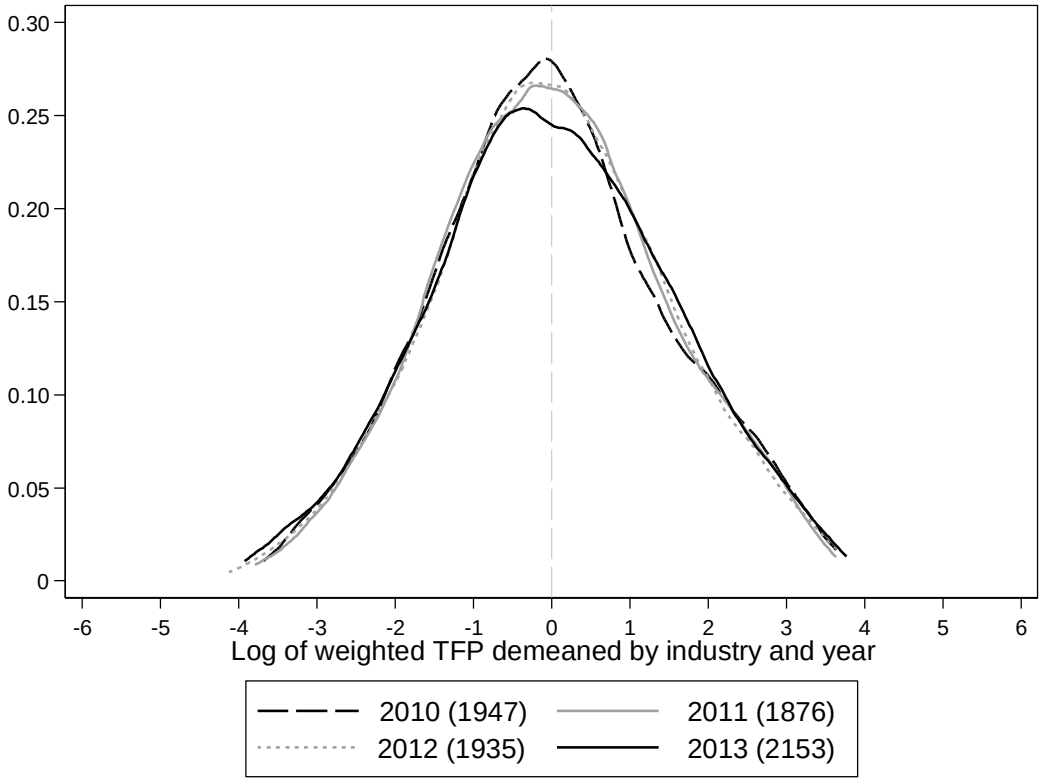
(m) 23: Other non-metallic mineral products



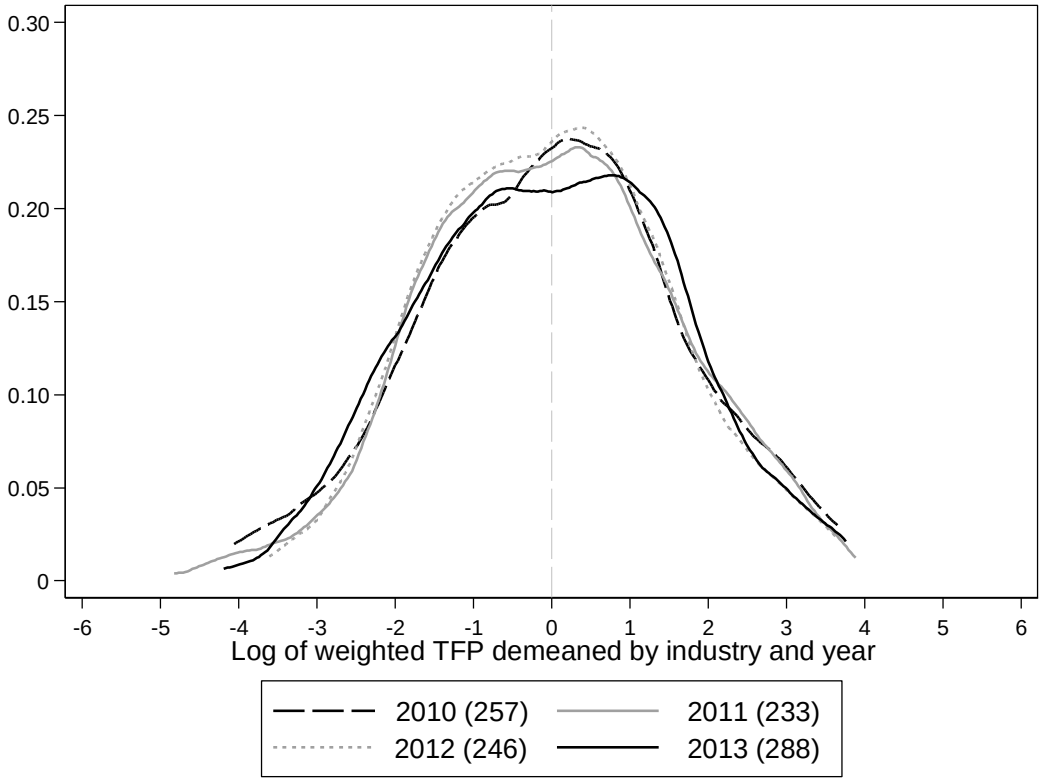
(n) 24: Basic metals



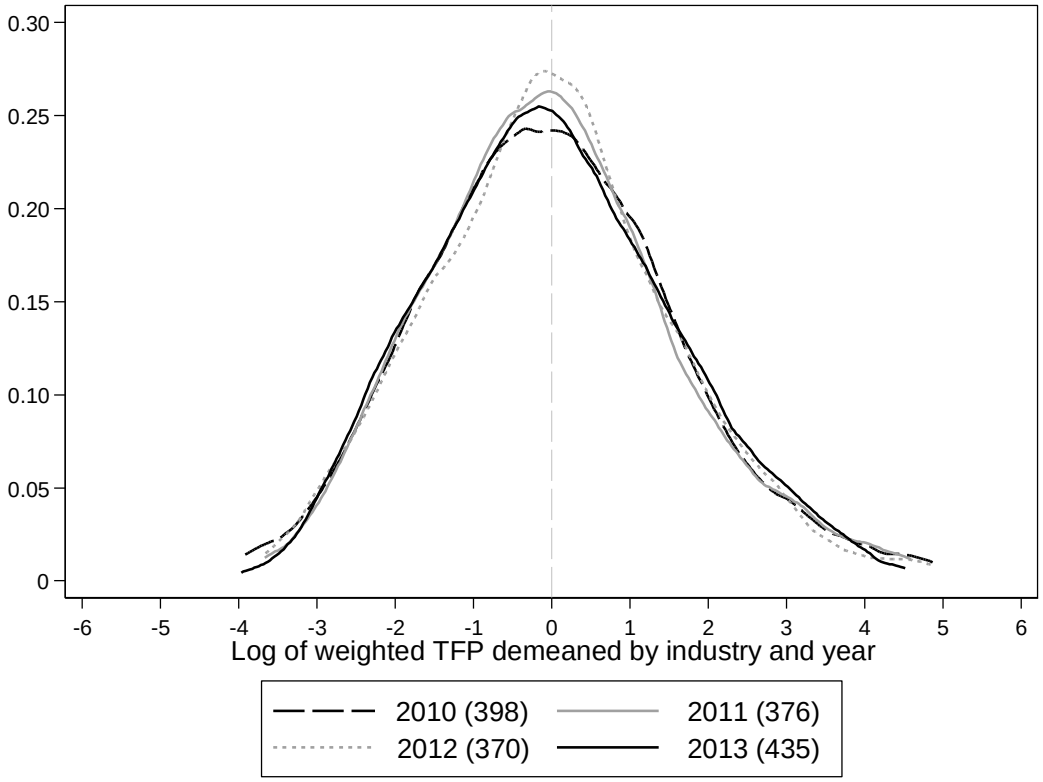
(o) 25: Fabricated metal products, except machinery and equipment



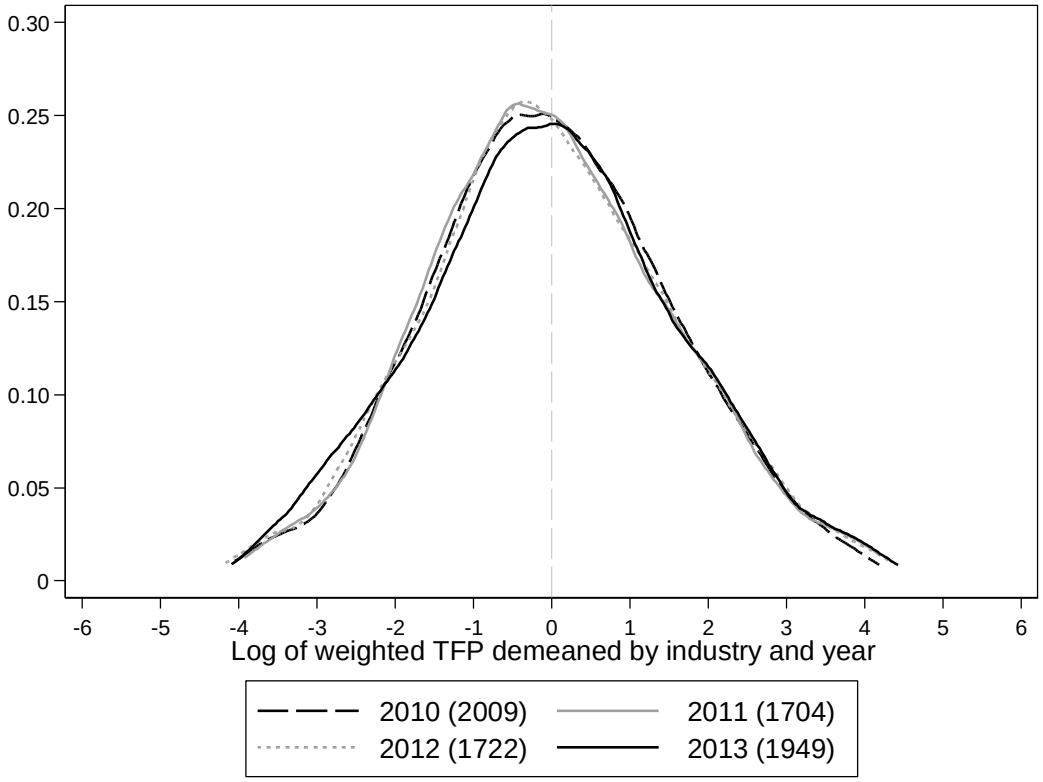
(p) 26: Computer, electronic, and optical products



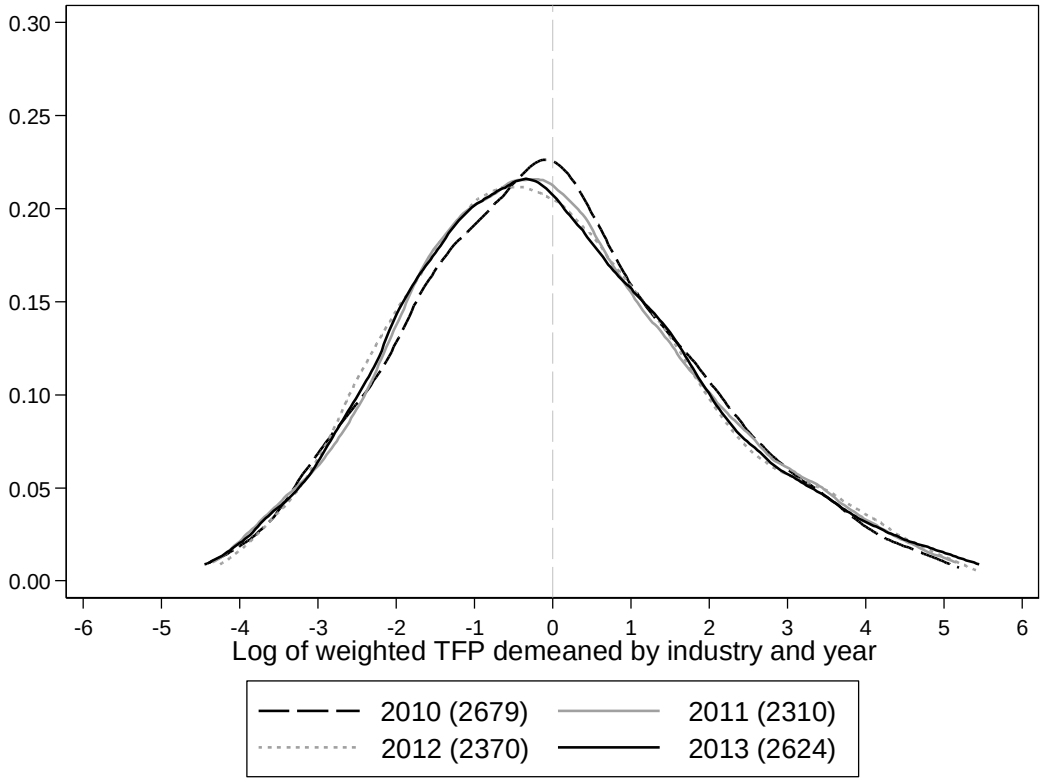
(q) 27: Electrical equipment



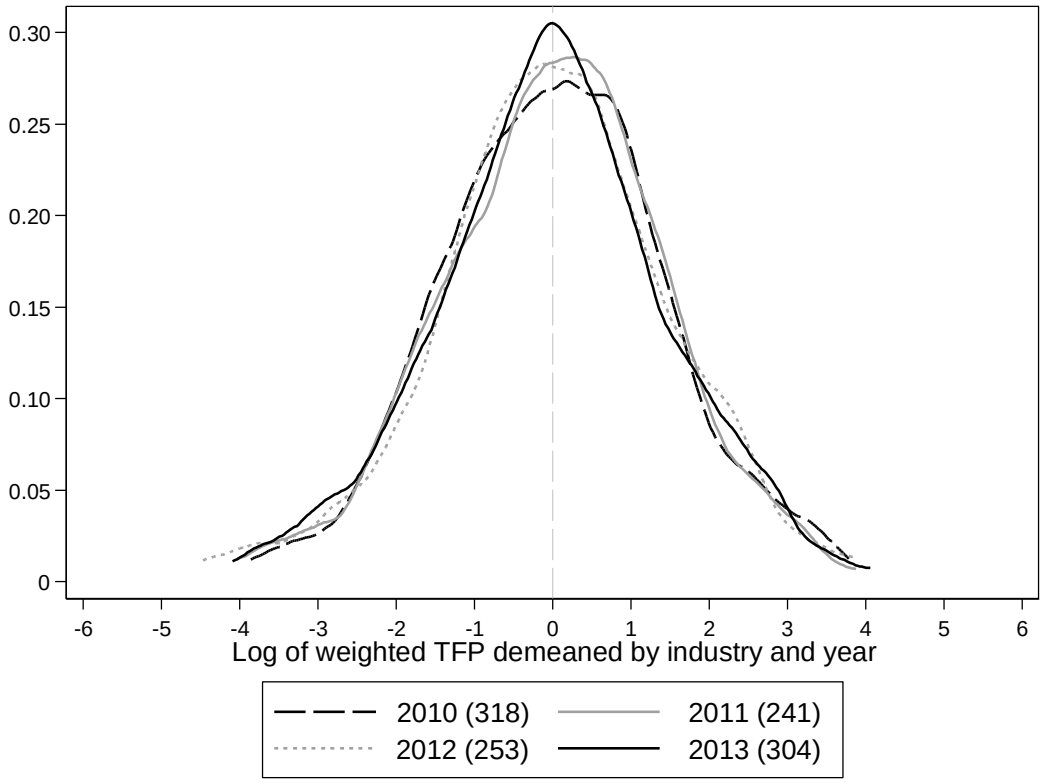
(r) 28: Machinery and equipment not elsewhere classified



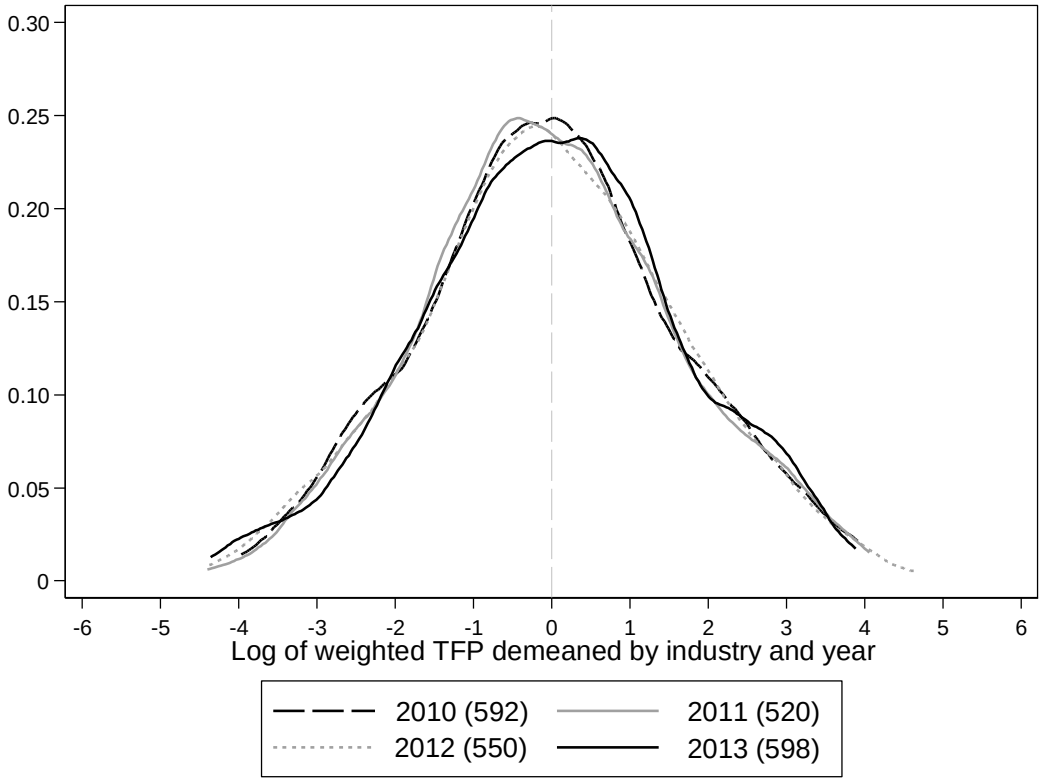
(s) 29: Motor vehicles, trailers, and semi-trailers



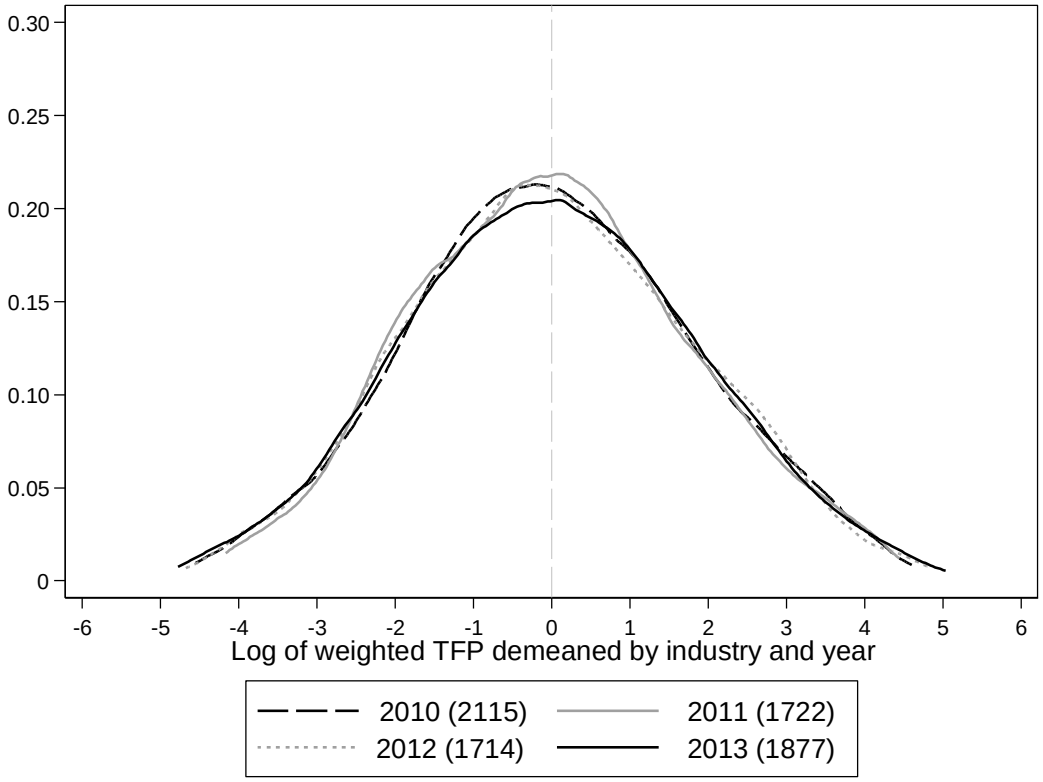
(t) 30: Other transport equipment



(u) 31: Furniture



(v) 32: Other manufacturing



Note: Refer to United Nations (2008: 63–7) for the exact contents of each industry.

Source: Authors' calculations based on TFP results from regressions on CIT–IRP5 data.