OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 81, 4 (2019) 0305–9049
doi: 10.1111/obes.12285

Productivity Growth, Human Capital and Technology Spillovers: Nonparametric Evidence for EU Regions*

Harald Badinger,† Peter H. Egger‡ and Maximilian von Ehrlich§

†Vienna University of Business Administration and Economics, Austrian Institute of Economic Research (WIFO) and CESifo, Althanstrasse 39-45, A-1090 Vienna, Austria (e-mail: harald.badinger@wu.ac.at)
‡ETH Zürich, CEPR, CESifo, Leverhulme Centre for Research on Globalisation and Economic Policy (GEP) at the University of Nottingham, and Oxford University Centre for Business Taxation (OUCBT), ETH Zürich, KOF, Weinbergstrasse 35, WEH E6 8092, Zürich, Switzerland (e-mail: egger@kof.ethz.ch)
§Department of Economics, University of Bern and CESifo, Schanzeneckstrasse 1, A211, 3001 Bern, Switzerland (e-mail: maximilian.vonehrlich@vwi.unibe.ch)

Abstract

This paper assesses the strength of productivity spillovers nonparametrically in a data set of 12 industries and 231 NUTS2 regions in 17 European Union member countries between 1992 and 2006. It devotes particular attention to measuring the catching up through spillovers depending on the technology gap of a unit to the industry leader and the local human capital endowment. We find evidence of a nonlinear relationship between the technology gap to the leader as well as human capital and growth in logs. Spillovers are smallest for units with a medium-high technology gap to the leader, especially for regions where human capital endowments are low.

I. Introduction

A large body of empirical work in macroeconomics emphasizes the role of total factor productivity (TFP) spillovers through knowledge diffusion for catching up and convergence. Nelson and Phelps (1966) suggested that the extent of knowledge spillovers depends on two factors, the distance to the technological frontier (the technology gap) and an economic unit's knowledge stock or human capital endowment. There is now broad evidence on the importance of either one of the two for catching up and convergence. Virtually all of this

JEL Classification numbers: C14; N10; N14; O33; O47.
*We thank Jonathan Temple and two anonymous reviewers for numerous comments on earlier versions of the manuscript. Moreover, the authors gratefully acknowledge comments by Wouter Vermeulen and participants at the annual conference of the International Institute of Public Finance at Dublin, 2015.

evidence assumes a parametric if not a (log-)linear relationship between spillovers and TFP growth. Little is known about the appropriateness of this assumption and the actual form of the relationship.

For instance, Benhabib and Spiegel (1994) identified two roles of human capital levels for economic growth in a large cross section of countries in 1965–85: for steady-state growth of TFP and for catching up, i.e. the absorption of spillovers from the technology leader. Griffith, Redding and Van Reenen (2004) assessed the determinants of TFP growth in a panel of OECD countries and manufacturing industries in 1974–90. Their findings suggest that human capital as well as R&D levels affect TFP growth and convergence. Kneller and Stevens (2006) found further support along those lines in a panel of industries and OECD countries in 1973–91, though suggesting that human capital was more robust a driver of spillovers than R&D in their data.[1] Cameron, Proudman and Redding (2005), using data on productivity in UK manufacturing industries, show that international trade is another key factor – beyond absorptive capacity – in facilitating technology transfer. In all of the just-mentioned work, a parametric relationship between catching up and human capital was assumed.

The goal of this paper is to apply nonparametric rather than parametric estimation techniques in assessing the functional form of the relationship between the technology gap and human capital for TFP growth and catching up. This is accomplished in a panel data set on 231 NUTS2 subnational regions (of 17 European Union member countries) and 12 industries over the period 1992–2006.[2] The findings are aligned with ones in earlier work to the extent that the speed of convergence is positively related to the size of the TFP gap *on average*, and that the speed of convergence towards the TFP leader is positively related to a unit's level of human capital *on average*. Yet, the nonparametric evidence reveals large areas in technology-gap and human-capital-endowment space where monotone convergence is absent and some regions with an intermediate gap to the leader do not benefit from the gap, e.g. due to an absence of absorptive capacity. Standard, monotone convergence in TFP growth to the industry leader as suggested by earlier, parametric work applies only to part of the region–industry units in the data.

The nonparametric estimator explains the lion's share of the variation in the data on TFP growth, while the explanatory power of the parametric estimator is relatively small. Moreover, the nonparametric estimator reveals a much greater variance in the marginal effect of human capital across regions and industries than the parametric estimator. Hence, allowing for flexible functional forms when assessing convergence processes and spillovers appears desirable and turns out to be qualitatively and quantitatively important. There are also systematic differences between the parametric and the nonparametric gradient functions of TFP growth with respect to the TFP gap to the industry leader and with respect to the human capital stock. In particular, the parametric estimator tends to overestimate the role of TFP gaps for regions with a medium-sized technology gap to the leader while it

[1] At the level of the firm, Griffith, Harrison and Van Reenen (2006) found evidence in support of R&D as a key determinant of catching up in TFP growth. Along different lines, the studies by Coe and Helpman (1995) and Coe, Helpman and Hoffmaister (1997) emphasized the role of international trade and foreign direct investment as transmission channels of international R&D spillovers (see Keller, 2004, for a survey of this and related work).

[2] *NUTS2* is a classification adopted by the statistical office of the European Union, Eurostat. It refers to regions of a size of 0.8–3 million inhabitants all over the European Union.

tends to underestimate it for relatively small and relatively large gaps. Overall, and unlike the parametric estimator, the nonparametric estimator suggests a varied pattern regarding the role of human capital for TFP growth and catching up.

The remainder of the paper is organized as follows. Section II introduces the data and measurement issues. Section III outlines a flexible nonparametric empirical model of productivity growth and discusses the results. Section IV concludes with a summary of the key findings.

II. Data

The empirical analysis in this paper involves two types of variables, one relating to TFP (a region's gap to the industry leader in an initial period as well as its average annual growth) and one relating to human capital endowments. Since TFP is not observed directly, we follow Griffith *et al.* (2004) for measurement.

Construction of TFP indices

Define $\Delta \ln Y_{it} \equiv \ln Y_{it} - \ln Y_{it-1}$ as the log change in region–industry pair i's value added in real terms between periods $t-1$ and t, $\Delta \ln L_{it} \equiv \ln L_{it} - \ln L_{it-1}$ as the log change in labour, $\Delta \ln K_{it} \equiv \ln K_{it} - \ln K_{it-1}$ as the log change in capital stock, and $\tilde{\alpha}_{it} \equiv 0.5(\alpha_{it} + \alpha_{it-1})$ as the average cost share of labour in value added in periods t and $t-1$. In particular, the cost share of labour is measured by the ratio of the wage sum and gross value added in region–industry pair. Then, the log change in i's TFP can be defined on the basis of a translog-production-technology-based superlative index (see Caves, Christensen and Diewert, 1982, p. 81) as[3]

$$\text{TFP growth}_{it} \equiv \Delta \ln A_{it} = \Delta \ln Y_{it} - \tilde{\alpha}_{it} \Delta \ln L_{it} - (1 - \tilde{\alpha}_{it}) \Delta \ln K_{it}. \tag{1}$$

Use $\ln \bar{V}_{it}$ to denote the geometric mean of a generic variable $\ln V_{it}$ within an industry and a year across all regions,[4] $D \ln V_{it} \equiv \ln V_{it} - \ln \bar{V}_{it}$ to denote i's deviation from $\ln \bar{V}_{it}$ in the same industry and year t, $\ln v_{it} \equiv D \ln V^*_{it} - D \ln V_{it}$ to denote the difference between the sector-year specific technology leader (*) and unit i in $D \ln V_{it}$, and $\sigma_{it} \equiv 0.5(\alpha_{it} + \bar{\alpha}_{it})$. Then, we may define

$$\text{TFP gap}_{it} \equiv \ln a_{it} = \ln y_{it} - \sigma_{it} \ln l_{it} - (1 - \sigma_{it}) \ln k_{it}. \tag{2}$$

[3] The translog production function represents a second-order approximation to a nonparametric production function. This is the main reason behind the large volume of academic work on technology and catching up in macroeconomics, international economics, and industrial organization which employs the superlative total-factor-productivity index based on Caves *et al.* (1982) as used here. An entirely different approach to determine technology in terms of a gap to the frontier is the one of nonparametric or semi-parametric stochastic-frontier analysis (see, e.g. Simar and Wilson, 2007). We resort to the former approach here, as the present paper is focused on the estimation of features of the technology gap to the leader by nonparametric methods rather than nonparametric estimation of that gap per se. From the viewpoint of inference, it would be relatively demanding to adopt a two-step approach based on fully nonparametric estimates in both steps, the one determining the technology gap and the one assessing the determinants of its growth.

[4] Note that $\ln \bar{V}_{it}$ carries an index i since i refers to region–industry pairs and the geometric mean is industry (-year)-specific.

Hence, information on the cost share of labour in value added, α_{it}, on value added in real terms, Y_{it}, on employment, L_{it}, and on the capital stock, K_{it} is required to measure TFP growth$_{it}$ and TFP gap$_{it}$.

Data sources

Information about α_{it}, Y_{it}, L_{it}, and K_{it} is based on data from Cambridge Econometrics. L_{it} and Y_{it} are measured directly with 2006 being the base year for the deflator. K_{it} is calculated by using the perpetual inventory method, using data on gross fixed capital formation, I_{it}, assuming a depreciation rate of 6%, $\delta = 0.06$, and an initial capital stock of $K_{i,1991} = \sum_{t=1980}^{1985} I_{it}$, so that $K_{it} = (1-\delta)K_{i,t-1} + I_{it}$ for all $t = 1992, \ldots, 2006$.[5] As in Harrigan (1997) and Griffith *et al.* (2004), we exploit the properties of the translog production function (which may be viewed as a second-order approximation to a nonparametric production technology) to smooth region-industry specific labour shares α_{it}, which are obtained as predicted values of a regression of the observed labour shares on a country-industry-specific fixed effect and the log of the capital–labour ratio, whose parameter is industry-specific. Data on regional human capital stocks H_{it} as one measure of absorptive capacity are based on the European Union's Labour Force Survey and the European Values Study. We employ information on the share of workers with at least secondary education which varies across NUTS2 regions, but not across industries. Overall, we have data for 10 industries across 231 regions and 15 years such that our estimates correspond to $n = 34,650$ observations.[6]

III. Empirical framework

Total factor productivity growth and convergence

A general, nonparametric catching-up process for TFP growth of the region–industry pair i at time t in the spirit of Benhabib and Spiegel (1994) and Griffith *et al.* (2004) may be formulated as

$$\Delta \ln A_{it} = m(X_{it}) + u_{it}, \text{ where } X_{it} = (H_{it-1}, \ln a_{it-1}, X_{it}^d), \tag{3}$$

where X_{it}^d is a vector of categorical variables in order to allow for fixed effects in three dimensions – regions, industry and time – along the lines of Racine and Li (2004).[7] Of course, the process in equation (3) is not consistent with unconditional convergence in general: there will be convergence to the same steady state only as long as $\Delta \ln A_{it}$ increases monotonically with the technology gap, $\ln a_{it-1}$, and as long as fixed effects are absent, which should not be expected to generally hold empirically; moreover, there will be uniform convergence to the region-industry-i-specific steady state in year t across the i and t only as

[5] Others such as Leamer (1984) or Harrigan (1999) assume a much higher rate of depreciation of around 15%. However, the results reported below are largely insensitive to using this value instead of one of 6%.

[6] Our data set covers the following industries following the NACE classification: Food, beverages and tobacco; Textiles and leather etc.; Coke, refined petroleum, nuclear fuel and chemicals etc.; Electrical and optical equipment; Transport equipment; Other manufacturing; Hotels and restaurants; Transport, storage and communications; Financial intermediation; Real estate, renting and business activities.

[7] See also the references in Li and Racine (2007) as well as section 11.3.2 in Henderson and Parmeter (2015) on this issue.

long as $\Delta \ln A_{it}$ increases monotonically with $\ln a_{it-1}$. This paper's main interest is to reveal the functional form of $m(X_{it})$, to contrast the findings with a parametric form as assumed in earlier work, and to outline conclusions for economic policy and future research. We acknowledge that human capital at the regional level is an endogenous variable, and thus we refrain from a causal interpretation of the effects.

Nonparametric estimation of technology spillovers

In this subsection, we are concerned with the specification of $m(X_{it})$ in equation (3). We specify the convergence forces in TFP as a potentially nonlinear function of the technology gap to the industry leader, $\ln a_{it-1}$, of human capital endowment, H_{it-1}, and of fixed effects captured by the categorical variables in X_{it}^d. For this, we employ a multivariate local linear estimator with an Epanechnikov product kernel for the two continuous variables and the three categorical variables in X_{it} with variable-specific bandwidths (see Racine and Li, 2004; Henderson and Parmeter, 2015). Let us use k to index one of the five columns in X_{it}. Moreover, let us index all data points (i.e. region–industry–time tuples) by j and denote the size of the data set by n. Then, the local linear regression model for all units j in the neighbourhood of unit s in equation (3) may be formalized as

$$\sum_{j=1}^{n} \{\Delta \ln A_j - \gamma_s - (X_j - X_s)\beta_s\}^2 \prod_{k=1}^{5} K_k \left(\frac{X_{k,j} - X_{k,s}}{b_k} \right), \tag{4}$$

where γ_s is a constant (which is specific to s) and the kernel for the k-th column of X_{it} involves a respective optimal bandwidth b_k. The prediction of $m(X_{it})$ is based on this smoother whose five optimal bandwidths are chosen from a leave-one-out cross-validation procedure. Confidence intervals of the local point estimates are computed by following Racine and Li (2004).[8]

We also estimate the gradients $\partial \widehat{\Delta \ln A}_{it} / \partial \ln a_{it-1}$ in H-$\ln a$-space by following Racine and Li (2004) and, specifically, Hayfield and Racine (2008). For illustration, we utilize three-dimensional plots of $\partial \widehat{\Delta \ln A}_{it} / \partial \ln a_{it-1}$ against $\ln a_{it-1}$ and H_{it-1}.

Results

In this section, we summarize the results of the nonparametric empirical analysis by way of plots. Three-dimensional plots generally come in two panels, one for the level function $\widehat{\Delta \ln A}_{it}$ (panel A) and one for the gradient function $\partial \widehat{\Delta \ln A}_{it} / \partial \ln a_{it-1}$ (panel B). In all three-dimensional plots we display transparent smoothed functions as well as point predictions using the following indicators to illustrate significance at the 5% level: crosses and circles for positive and negative values, respectively, and black and grey for statistically significant (at 5%) and insignificant values, respectively.

[8] Chu, Henderson and Parmeter (2017) propose an alternative estimator for discrete variables. While the estimator of Racine and Li (2004) has a tendency to undersmooth (i.e. select a very small bandwidth with discrete, unordered regressors) in some data situations, the one of Chu, Henderson, and Parmeter is less prone to that problem. However, their simulations suggest that the performance of the estimator of Racine and Li (2004) is often similar to theirs. We will report on the selected bandwidths in the following subsection.

The nonparametric estimation results illustrated in Figure 1a,b can be summarized as follows. First, the estimated levels function in Figure 1a is not monotonic. Second, having a small or a very large technology gap to the leader is best for convergence in general and independent of human capital endowments. However, with a medium-sized technology gap, the convergence to the leader is faster, on average, with a high level of human capital as the technology gap rises.

The relative importance of a technology gap to the industry leader and of human capital for convergence can be visualized by plotting the fractions of units at certain levels of a technology gap or human capital where the point estimate for the gradient is positive and significant. The corresponding fractions for $\partial\widehat{\Delta \ln A_{it}}/\partial \ln a_{it-1}$ across different levels of $\ln a_{it-1}$ and H_{it-1} are displayed in Figure 2a,b. Again, a positive, statistically significant TFP-growth gradient is more likely at either a quite small or a very large gap to the leader. The fraction of positive significant values of $\partial\widehat{\Delta \ln A_{it}}/\partial \ln a_{it-1}$ is less clearly related to human capital endowments according to Figure 2b. However, as indicated above, a higher level of human capital helps regions with a medium-sized technology gap to the leader to catch up faster and more systematically. The latter is not obvious from Figure 2b, however.[9]

An analysis of variance of an indicator variable which is unity for observations, which catch up with the leader in a significantly positive way, and zero else, using region- and industry-specific indicators reveals the following. First, only about one-twelfth of the variation in this variable is explained by industry-specific and region-specific indicators, with the the industry dimension being almost twice as important as the regional dimension. Overall, there is relatively little concentration of TFP spillovers across specific industries and even less so across specific regions. The fraction of significantly positive spillovers varies across industries with values between 0.45 (Transport equipment) and 0.79 (Other manufacturing) being centred around an average of 0.63 and across regions with values between 0.41 (Dél-Dunántúl in Hungary) and 0.80 (South Yorkshire in the United Kingdom).

The nonparametric estimates reveal more variation in the role of H_{it} compared to our version of a parametric model in the spirit of Griffith *et al.* (2004). In the parametric model, $\Delta \ln A_{it}$ is (log-)additive in a constant, H_{it-1}, $\ln a_{it-1}$, and an interaction $H_{it-1} \ln a_{it-1}$. While such a parametric approach yields for our sample an average marginal effect of about -0.043 for H_{it-1}, the fully nonparametric approach predicts one of about 0.058 (i.e. the sign is different). The standard deviation of the average marginal effect of H_{it-1} is 0.023 and 0.034 in the parametric and the nonparametric model, respectively.

The parametric approach (including the region-, sector- and time-fixed effects) explains about 11.5% of the variation in $\Delta \ln A_{it}$, while the fully nonparametric framework explains about 99.9%. The latter could be an outcome of undersmoothing in the fixed-effects dimensions. Regarding the latter, let us add the following. First of all, the unordered, discrete variables capturing the year, sector, and regional indicators contain 15, 12, and 231 categories, respectively. The respective kernel function for the kth categorical variable depends

[9] Notice that the conditional variance on the three dimensions of fixed effects in the human-capital-endowment variable is small relative to the corresponding unconditional variance. The reason is that human capital endowments change relatively sluggishly over short time periods. In a model without region-specific fixed effects, the role of human capital would turn out more clearly positive for catching up and convergence.

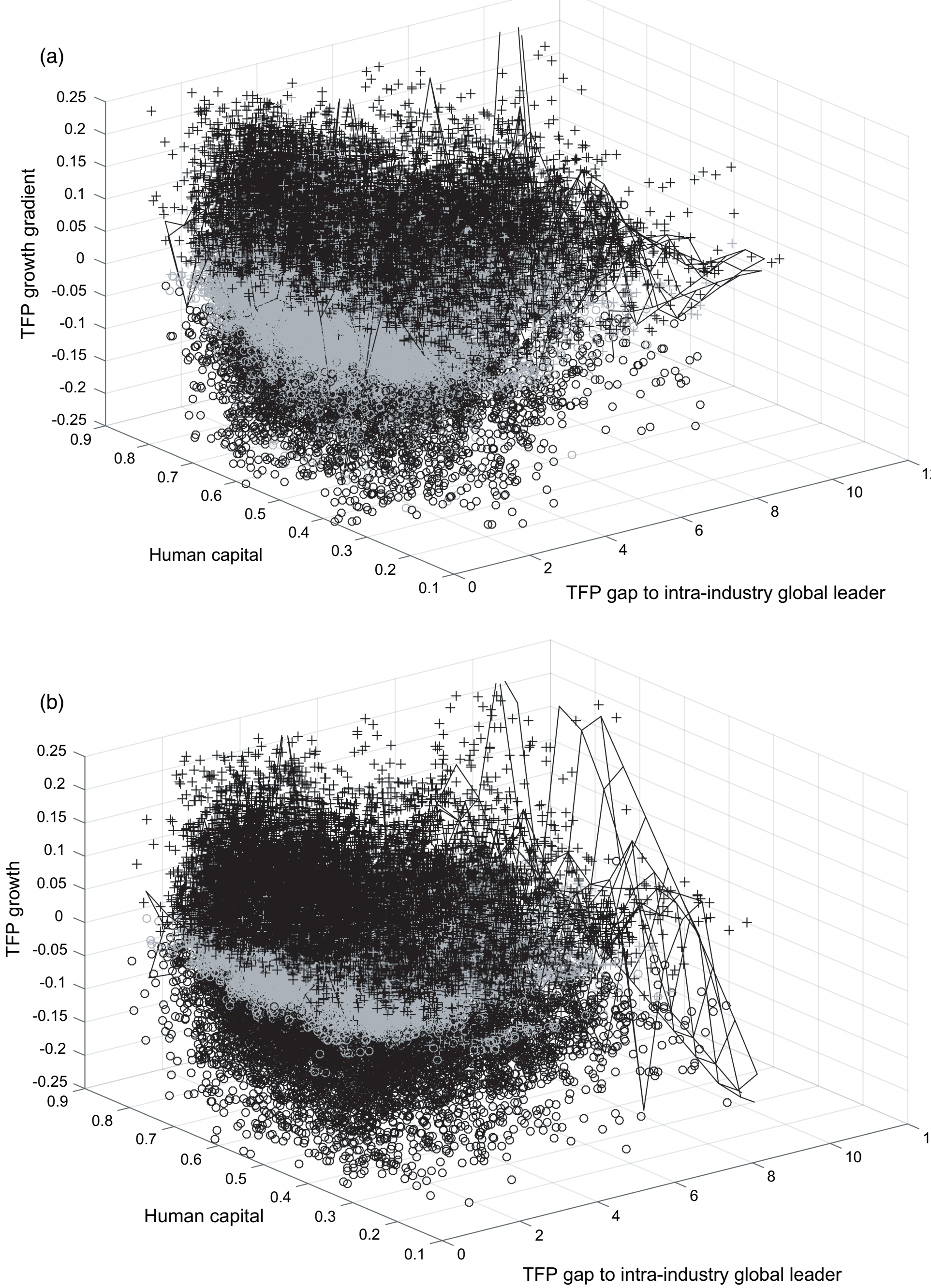

Figure 1. Nonparametric spillovers
Notes: The estimates are based on a sample of 34,650 observations. Panel (a) illustrates the estimates for the level $\Delta \ln A_{it}$ where we use the following indicators: '+' and 'o' for positive and negative values, respectively, and black versus grey colours for statistically significant (at 5% level) and insignificant values, respectively. Panel (b) refers to the estimates for the gradient $\partial\widehat{\Delta \ln A_{it}}/\partial \ln a_{it-1}$, using the same colouring as in panel (a).

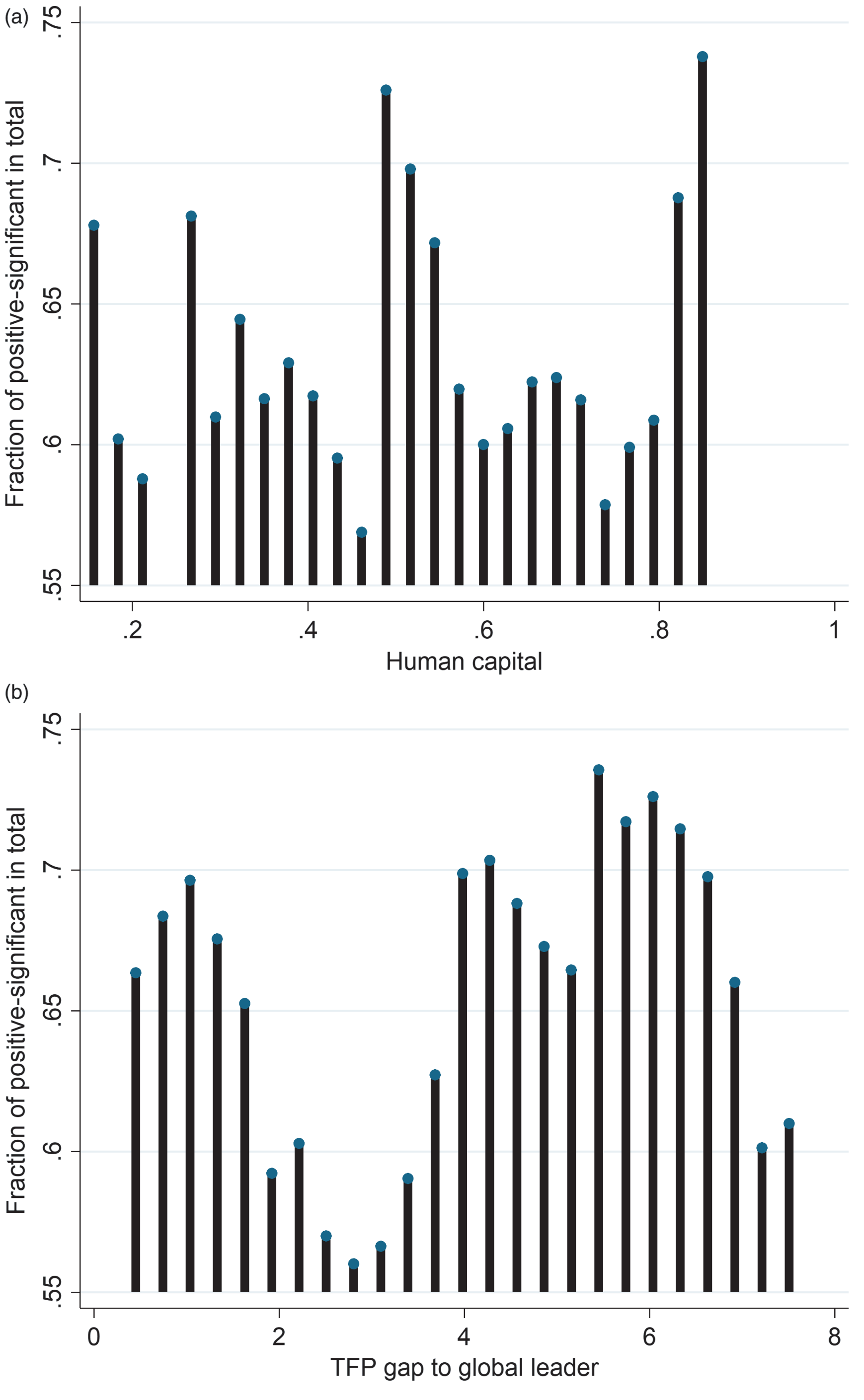

Figure 2. Positive and significant nonparametric spillovers
Notes: Panel (a) plots the fractions of observations within 25 equally sized bins of $\mathfrak{D} \ln A_{it-1}$ for which the nonparametric estimator predicts a significantly positive gradient $\partial \widehat{\Delta \ln A_{it}} / \partial \ln a_{it-1}$. Panel (b) plots these fractions against 25 equally sized bins of H_{it-1}.

on the smoothing parameter λ_k. The latter is bounded between 0 and $N_k/(N_k-1)$, where N_k is the number of categories in variable k. With $\lambda_k=0$ there is no smoothing of the fixed effects in the respective dimension, while with $\lambda_k=N_k/(N_k-1)$ there is complete smoothing. Hence, the closer λ_k is to zero, the larger is the danger of undersmoothing. The latter should be particularly an issue with categorical variables pertaining to regional effects rather than years or sectors, as the number of regional categories is much larger than the ones of years and sectors are. With the data at hand, $\lambda_k=0.0039$ for years, $\lambda_k=0.0025$ for sectors, and $\lambda_k=0.4653$ for regions. Hence, there is relatively strong smoothing across regions but less so for years and sectors.

In Figure 3a,b we provide a comparison of the nonparametric and parametric functions of the TFP-growth levels and gradients, subtracting the nonparametric from the parametric estimates.[10] In those figures we use the same indicators as above for the signs and statistical significance of estimates.

A close inspection of Figure 3a,b suggests the following. First, the 95%-confidence intervals of the model predictions of $\Delta\ln A_{it}$ are overlapping between the parametric and the nonparametric estimates for only about 21% of the observations. Moreover, the 95%-confidence intervals of the model predictions of the gradient $\partial\widehat{\Delta\ln A_{it}}/\partial\ln a_{it-1}$ are overlapping between the parametric and the nonparametric estimates for about 25% of the observations. Hence, there is a large degree of stochastic non-overlap between the parametric and the nonparametric model predictions. Second, in large subspaces in the $\ln a_{it-1}$-H_{it-1}-domain of the paper, namely for about 53% of the observations, the nonparametric function predicts higher levels of TFP growth – consistent with the data – than the parametric one. We obtain negative deviations, where the parametric estimator predicts higher levels of TFP growth than the nonparametric one. The latter occurs mainly for medium-low levels of TFP gaps. Third, the gradient $\partial\widehat{\Delta\ln A_{it}}/\partial\ln a_{it-1}$ tends to be biased downwards in the parametric specifications for about 59% of the observations, while it tends to be systematically upward biased for about 16% of the observations. The latter is mainly the case for regions with high levels of human capital endowments. Moreover, following Hsiao, Li and Racine (2007) we performed a robust consistent nonparametric test about whether the parametric model is correctly specified. In particular, we test the null hypothesis that the difference between the data and the parametric conditional expectation is zero almost everywhere. The importance of a nonparametric model is confirmed by the test which yields a P-value of 2.22×10^{-16} and thus clearly rejects the parametric specification. Note that this holds true for a number alternative parametric specifications we tested.

[10]The parametric predictions of the TFP-growth levels and gradient functions, respectively, are

$$\widehat{\Delta\ln A_{it}}=-0.030H_{it-1}+0.011\ln a_{it-1}-0.005H_{it-1}\ln a_{it-1}+\text{fixed effects}_{it},$$
$$\frac{\partial\widehat{\Delta\ln A_{it}}}{\partial\ln a_{it-1}}=0.011-0.005H_{it-1},$$

where fixed effects$_{it}$ is a compact representation of the fixed region, sector, and time effects.

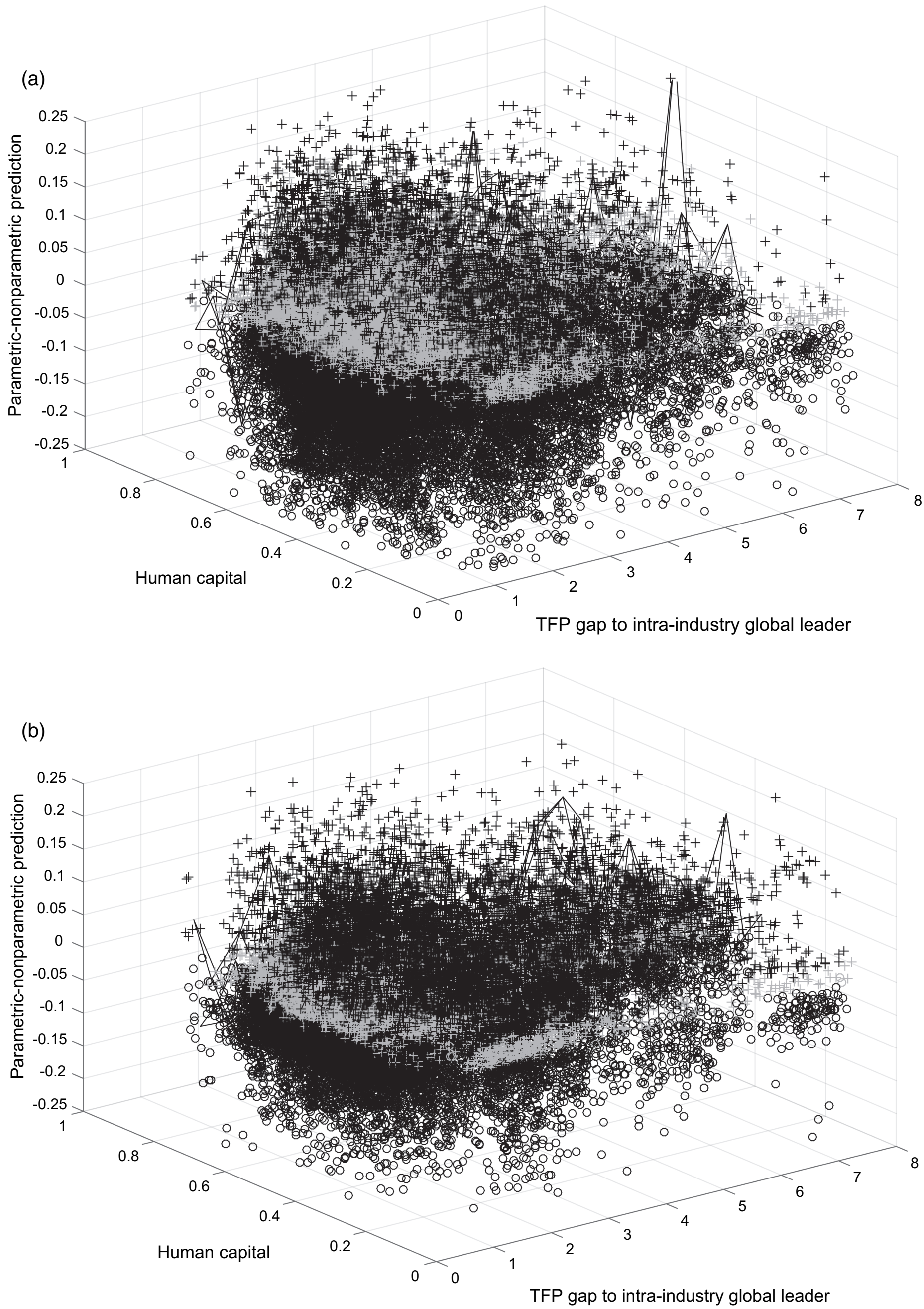

Figure 3. Parametric vs. nonparametric model predictions
Notes: The estimates are based on a sample of 34,650 observations. Panel (a) illustrates the differences of the parametric and nonparametric model estimates (parametric minus nonparametric) for the level $\Delta \ln A_{it}$ where we use the following indicators: '+' and 'o' for positive and negative values, respectively, and black versus grey colours for statistically significant (at 5% level) and insignificant values, respectively. Panel (b) refers to the differences of the parametric and nonparametric model estimates (parametric minus nonparametric) for the gradient $\partial \widehat{\Delta \ln A_{it}} / \partial \ln a_{it-1}$, using the same colouring as in panel (a).

IV. Conclusions

This paper studies the role of the technology gap and absorptive capacity of regions and industries for catching up. The functional relationship between TFP growth and the technology gap and human capital endowments features considerable nonlinearities and even non-monotonicities that can typically not be captured by parametric specifications. The estimates suggest that spillover effects from the technological leader are strongest to regions within an industry where the technology gap is either quite small or sufficiently large, at least in Europe. For a medium-sized technology gap, we do not identify positive spillovers. This provides implicit evidence of high levels of absorptive capacity at small technology gaps to the industry leader and a strong pull effect of technologically leading regions on ones with a very large technology gap. The nonparametric estimation reveals a much bigger variation in the role of human capital (absorptive capacity) for growth and convergence than the conventional, parametric approach. This appears particularly important when thinking of returns to human capital across regions and industries and the funding of education in the presence of financial constraints.

Final Manuscript Received: September 2018

References

Benhabib, J. and Spiegel, M. M. (1994). 'The role of human capital in economic development – evidence from aggregate cross-country data', *Journal of Monetary Economics*, Vol. 34, pp. 143–173.

Cameron, G., Proudman, J. and Redding, S. (2005). 'Technological convergence, R&D, trade, and productivity growth', *European Economic Review*, Vol. 49, pp. 775–807.

Caves, D., Christensen, L. and Diewert, E. (1982). 'The economic theory of index numbers and the measurement of input, output and productivity', *Econometrica*, Vol. 50, pp. 1393–1414.

Chu, C.-Y., Henderson, D. J. and Parmeter, C. F. (2017). On discrete Epanechnikov kernel functions. *Computational Statistics and Data Analysis*, Vol. 116, pp. 79–105.

Coe, D. T. and Helpman, E. (1995). 'International R&D spillovers', *European Economic Review*, Vol. 39, pp. 859–887.

Coe, D. T., Helpman, E. and Hoffmaister A. W. (1997). 'North-South R&D spillovers', *Economic Journal*, Vol. 107, pp. 134–49.

Griffith, R., Redding, S. and Van Reenen, J. (2004). 'Mapping the two faces of R&D: productivity growth in a panel of OECD industries', *Review of Economics and Statistics*, Vol. 86, pp. 883–895.

Griffith, R., Harrison, R. and Van Reenen, J. (2006). 'How special is the special relationship? Using the impact of U.S. R&D spillovers on U.K. firms as a test of technology sourcing', *American Economic Review*, Vol. 96, pp. 1859–1875.

Harrigan, J. (1997). 'Technology, factor supplies, and international specialisation', *American Economic Review*, Vol. 87, pp. 475–494.

Harrigan, J. (1999). 'Estimation of cross-country differences in production functions', *Journal of International Economics*, Vol. 47, pp. 267–293.

Hayfield T. and Racine, J. S. (2008). 'Nonparametric Econometrics: The np Package', *Journal of Statistical Software*, Vol. 27, pp. 1–32.

Henderson, D. J. and Parmeter, C. F. (2015). *Applied Nonparametric Econometrics*, Cambridge University Press, Cambridge, UK.

Hsiao, C., Li, Q., Racine, J. S. (2007). 'A consistent model specification test with mixed categorical and continuous data', *Journal of Econometrics*, Vol. 140, pp. 802826.

Keller, W. (2004). 'International technology diffusion', *Journal of Economic Literature*, Vol. 42, pp. 752–782.

Kneller, R. and Stevens, P. A. (2006). 'Frontier technology and absorptive capacity: evidence from OECD manufacturing industries', *Oxford Bulletin of Economics and Statistics*, Vol. 68, pp. 1–21.

Leamer, E. E. (1984). *Sources of International Comparative Advantage: Theory and Evidence*, MIT Press, Cambridge, MA.

Li, Q. and Racine, J. S. (2007). *Nonparametric Econometrics: Theory and Practice*, Princeton University Press, Princeton, NJ.

Nelson, R. and Phelps, E. (1966). 'Investment in humans, technological diffusion, and economic growth', *American Economic Review: Papers and Proceedings*, Vol. 61, pp. 69–75.

Racine, J. S. and Li, Q. (2004). 'Nonparametric estimation of regression functions with both categorical and continuous data', *Journal of Econometrics*, Vol. 119, pp. 99–130.

Simar, L. and Wilson, P. (2007). 'Estimation and inference in two-stage, semi-parametric models of production processes', *Journal of Econometrics*, Vol. 136, pp. 31–64.