Productivity Growth and the Speed of Convergence of Domestic Firms

Carlo Altomonte*   Enrico Pennings†

January 2008

Abstract

We investigate productivity convergence of domestic firms in a transition economy, Romania. In estimating total factor productivity we allow for varying returns to scale and control for both the endogeneity of the productivity shock and the omitted price variable bias linked to heterogeneous firms' market power. Consistently with our priors, we find that without controlling for the omitted price variable bias absolute convergence estimates are biased upwards. In terms of conditional convergence, we find that the speed of convergence across firms depends mainly on technology transfers from the frontier and, less markedly, by a number of regional and industrial characteristics such as the distance to the capital region, the minimum efficient scale and the absorptive capacity.

JEL classification: F23; L10; P20

Keywords: multinational firms, productivity, transition economies.

*Bocconi University, Milan.

†Erasmus University and Tinbergen Institute, Rotterdam. *Corresponding author:* Faculty of Economics H13-31, Erasmus University Rotterdam, Burg. Oudlaan 50, 3062PA Rotterdam, Netherlands. Phone +31 10 4082166. Fax +31 10 4089149. pennings@few.eur.nl

1 Introduction

Standard neoclassical economic theory suggests that, under diminishing returns and free movement of factors, per capita income levels within an economic area should converge over time to the same steady state value (Barro and Sala i Martin, 1991). However, such a view has been challenged since long by many authors[1], who have found a persistence of income disparities, arguing therefore that the pattern of cross-country growth is more consistent with endogenous growth, rather than neoclassical theories. Most of this early empirical literature has generally used either cross-section or time-series techniques. Islam (1995) advocates instead panel data approaches to estimate productivity growth convergence since, by incorporating country fixed effects, these models account for initial efficiency and thus test for conditional convergence. Lee, Pesaran and Smith (1997) comment that heterogeneity in speed of convergence from such a panel regression may bias the results. Bernard and Jones (1996) used both cross-section and time-series approach to measure the convergence of sectoral productivity in different industries with respect to aggregate productivity in a panel of fourteen OECD countries, and found no sign of convergence in manufacturing industries, but a different response of services. They also discussed the relevance of different concepts of convergence (β vs. σ-convergence)[2] as well as the importance of properly measuring productivity in order to obtain unbiased results.

Recently, the increasing availability of disaggregated cross-country data has revamped an interest in explaining differences in the sources and speed of convergence. By regressing GDP growth on the interaction of lagged GDP and an indicator of financial development, Aghion, Howitt and Mayer-Foulkes (2005) show the positive effect of financial development on the speed of convergence. Other recent studies have looked at TFP growth instead, using the interaction between the distance to the technology frontier and a variable for the speed of convergence as explanatory variable for growth. Using the Penn World Tables, Benhabib and Spiegel (2005) explain differences in the cross-country speed of convergence through schooling. In a panel of 12 countries over the period 1974-1990, Griffith, Redding and Van Reenen (2004) use in addition absorptive capacity and imports as determinants for the speed of convergence. Schooling and absorptive capacity appear to positively affect the speed of convergence, while imports do not have a significant impact. Employing a panel of 14 UK manufacturing industries, Cameron, Proudman and Redding (2005) do find that international trade significantly enhances the speed of technology transfer.

This paper examines a panel of some 48,000 firms operating in Romanian regions over the period 1996-2001. Romania represents a very interesting 'natural experiment' for our purposes since, before the start of transition from plan to market in 1995, the country experienced limited factor movements across its regions, associated to low regional disparities. After 1995, i.e. since when we have census data, disparities started to increase along the transition process, thus

[1] See Temple (1999) for a general overview of the empirical growth literature or Mohnen (1996) for a more specific survey of TFP growth.

[2] The concept of σ-convergence deals with the dispersion of productivity over time; β-convergence refers to a negative correlation between the initial level of productivity and its rate of growth. The latter is a necessary condition for σ-convergence, but not a sufficient one.

providing us with an ideal control for initial conditions. The paper makes several contributions to the convergence literature.

First, from a methodological point of view, our regional and industry specific TFP estimates derive from an average of firm-specific TFP estimated using the semi-parametric method by Levinsohn and Petrin (2003). The literature has typically used the labor share in value added as coefficient in the production function (instead of estimating this coefficient through a regression) and has assumed constant returns to scale[3]. The semi-parametric estimation of the production function allows instead for varying returns to scale and produces better TFP estimates than conventional methods such as OLS estimation of the production function, as the method controls for the simultaneity bias which arises when inputs and the error term are correlated. In addition, by using industry averages of a total of 10,650 firm specific estimates of TFP, measurement error is reduced, resulting in lower standard errors of the estimated coefficients. Furthermore, we explicitely discuss another typical problem of TFP estimation, and namely the omitted price variable bias induced by the correlation between individual firms' prices and their used inputs (see Klette and Griliches, 1996). We show how, failing to take this bias into account, the resulting estimated speed of convergence could be upward biased, and we discuss two methods to control for that.

Second, in contrast to the recent literature on the speed of convergence, this paper takes a more regional and sectoral perspective allowing to test new explanations for the speed of convergence. Sectoral and regional data on FDI inflows allow for a more direct test of the significance of openness on the speed of convergence with respect to country specific trade data. As an industry specific variable that can affect the speed of convergence, we employ the minimum efficient scale (MES). MES may affect the speed of convergence, as firms in industries where firms are on average larger are more likely to possess a sufficient level of absorptive capacity (Aitken and Harrison, 1999). Another industrial variable that is related to absorptive capacity and may affect the speed of convergence is the average ratio of intangibles to total assets, also employed in the analysis. Finally, we use the distance from the region to the capital as a region-specific variable that may have an impact on the speed of convergence.

As far as our results are concerned, we show that, in line with our theoretical priors, failing to take into account the omitted price variable bias in TFP measures results in a significant (more than 60%) upward bias in the detected speed of absolute convergence. In terms of conditional convergence, we employ an equilibrium correction model, as in Cameron et al. (2005), to show that domestic firms seem to benefit from technology transfers from the frontier, with the firms lying further from the technological frontier having a higher rate of TFP growth, in line with convergence theory. In terms of other covariates affecting convergence, the presence of FDI in horizontal, backward or forward industries does not seem to have a robust effect on the TFP convergence of domestic firms once distance from the frontier is taken into account, a result which sheds some new light on the spillover literature (e.g. Smarzynska Javorcik, 2004). We

[3]Bernard and Jones (1996) openly discuss the restrictions imposed by the measure of labor productivity when performing a convergence analysis, and thus estimate various specifications of TFP. They however perform their analysis at the aggregate country/industry level, still imposing constant returns to scale.

also find that a higher MES tends to stimulate internal gains in TFP by firms, as argued by Aitken and Harrison (1999). In terms of rate of technology transfers from the frontier, the latter tend to decrease in line with the the average distance of each firm from the capital region, characterized by a significantly higher level of per capita GDP. Finally, we do not find evidence of σ-convergence or divergence[4] when all firms are included in the analysis, with only a slight σ-divergence detected for the TFP of domestic firms.

The paper is organized as follows. Section 2 presents our dataset. In Section 3 we discuss the empirical strategy, whose results are outlined in Section 4. Section 5 concludes the paper.

2 The Romanian dataset

To analyze the micro sources of regional convergence in Romania, we employ a dataset composed of domestic firms and affiliates of foreign multinational enterprises (MNEs) operating for the period 1996-2001 in the manufacturing and construction industries, as retrieved from AMADEUS. The latter is a dataset provided by a consulting firm, Bureau van Dijck, containing balance sheet data in time series for a sample of roughly 7,000,000 companies operating in various European countries. In the case of Romania, the dataset covers the entire census of operating firms, since it reports the information recorded by the Romanian Chamber of Commerce and Industry, the institution to which all firms have to be legally registered and report their balance sheet data. In particular, we have retrieved information on the location of each firm within each of the eight Romanian regions, the industry in which these firms operate (at the NACE-4 level), as well as yearly balance sheet data on tangible and intangible fixed assets, total assets, number of employees, material costs, value of production and value-added.

The dataset retrieved from the Romanian census is analyzed in Table 1, and consists of 39,799 active firms at the beginning of the period (of which 36,634 are domestically owned and 3,165 display a multinational participation), then becoming 48,718 in 2001 (of which 41,981 domestic and 6,737 MNEs). These figures correspond to 95 per cent of all official firms operating in Romania in manufacturing and construction, with the exception of 2001, where this percentage drops to 85 per cent. Entry rates tend to overcome the exit of firms at the beginning of the period, while exit rates grow larger towards the end, a dynamic not surprising for a transition country, where soft budget constraints are progressively removed. Moreover, the share of multinational enterprises increases from 8 to 14 per cent of the total. For both the domestic and multinational firms, the food (NACE-15) and construction (NACE-45) industries are the two largest in terms of number of entities over the considered time span.

[Table 1 about here]

The sample coverage is lower if we consider only those firms for which information is available for all the variables of interest in the calculus of TFP with the latter restricted sample covering

[4]To calculate σ-convergence we regress the yearly standard deviation of firms' productivity in each industry / region over a time trend.

around 50 per cent of all official firms.

Given the microfounded nature of the data, we have to address three methodological issues. First, the panel data retrieved from AMADEUS is unbalanced, i.e. it incorporates firms' entry and exit, which have to be properly controlled for. Second, information on the ownership structure is not available for all firms. Third, and most importantly, the sample has to be validated, i.e. when aggregated firm-level observations should be able to reproduce fairly well the evolution of regional dynamics at the country level.

The first issue arising from our data is related to the treatment of firms' entry and exit. To this extent, the year in which the first observation is recorded denotes a firm's entry, while exit is assumed to take place in the year after which no new information is available in the dataset. Both our entry and exit rates so calculated are in line with the ones reported from official statistics for Romania (data available from the Romanian Chamber of Commerce).

Second, we have included in the sample only those firms for which information on the ownership structure is available: in particular, a firm is considered as foreign MNE affiliate if more than 10 per cent of its capital is foreign owned, and domestic otherwise[5]. Clearly, given the nature of our data, it could be the case that a firm exits and then reappers under a new name, eventually due to a change in ownership. In order to gauge the magnitude of this issue, we have compared different yearly releases of AMADEUS, finding that, given a MNE in year 2000, there is a 15 per cent chance that the same firm is a domestic one before that year, while the probability of the opposite event (a firm switching from MNE to domestic) is negligible[6].

None of these data issues is however critical for our exercise, since the aim is to derive a correct measure of the average productivity for a sample of domestic firms. If we incorrectly attribute the multinational status to that 15 per cent of firms which sometime before 2001 were still domestic, we de facto exclude them from our dependent variable (domestic firms' TFP). The latter exclusion leads to a more conservative TFP measure, if we assume that MNEs acquire the most productive domestic firms. A similar conservative outcome derives from a possible measurement error in the entry and exit rates: if in the Romanian transition to the market economy there are soft budget constraints, so that firms tend to survive also when they should not, then TFP in our sample is relatively less influenced by the selection effect driving out inefficient domestic firms.

The only relevant bias might actually derive from an unbalanced territorial distribution of the dataset: if the selection of firms according to data availability generates in our sample an over-representation of regions which, for some unobserved reasons, tend to be subject to relatively larger productivity shocks, we might observe a spurious correlation between the latter and the presence of MNEs. To this extent, we have retrieved from our restricted sample (the one we actually use for TFP estimation) a yearly measure of regional output, summing the individual firms' revenues operating in each region. We have then correlated these figures with the official

[5]The implications of a varying degree of foreign ownership in MNEs' affiliates for Romania are discussed by Spatareanu and Smarzynska Javorcik (2006).

[6]Due to the limited coverage of earlier versions of the dataset, we have been able to identify only a sub-sample of firms for which it is possible to track the entire ownership history for the period 1997-2000.

regional figures for Romania, obtaining a significant positive correlation of 0.87[7]. As a result our firm-level data seem to belong to an unbiased sample, being able to reproduce the actual evolution of output in Romania.

3 Methodology

One of the main challenges in the productivity/convergence literature is related to the accurateness with which TFP growth and its relative levels are measured. The literature usually opts between two main approaches: the superlative index number approach and production function estimation (Griffith, Redding and Simpson, 2006). An advantage of index numbers is that they allow for a more flexible form in the production function, typically a translog. However, the key assumptions behind the superlative index number measures are constant returns to scale and perfect competition, two features which seem very restrictive in the case of a transition country such as Romania.

We have thus opted for the calculation of TFP as the Solow residual of an estimated firm-specific production function (Cobb-Douglas), where no a priori assumption is imposed on the industry-specific returns to scale. In particular, in order to calculate firm-specific productivity, we have initially followed the standard approach of deflating our balance sheet data using disaggregated industry price indexes[8]. We have proxied output with deflated sales, given the better quality of these time series with respect to the ones reporting value added. The number of employees has been used as a proxy for the labour input, and the deflated value of tangible fixed assets as a proxy for capital. We have then estimated within each industry semi-parametric productivity measures at the firm level[9]. In fact, using ordinary least squares when estimating productivity implies treating labor and other inputs as exogenous variables. However, as pointed out by Griliches and Mareisse (1995), profit-maximizing firms can immediately adjust their inputs (in particular capital) each time they observe a productivity shock, which makes input levels correlated with the same shocks. Since productivity shocks are unobserved to the econometrician, they enter in the error term of the regression. Hence, inputs turn out to be correlated with the error term of the regression, and OLS estimates of production functions suffer from the so-called simultaneity bias. Olley and Pakes (1996) and Levinsohn and Petrin (2003), henceforth OP and LP, have developed two similar semi-parametric estimation procedures to overcome this problem using, respectively, investment and material costs as instruments for the

[7]Since our sample does not include all NACE industries (in particular agriculture and services), we have subtracted from official regional GVA data the output of those industries not present in our dataset. The correlation between our sample and the official regional data comprising all NACE industries is instead 0.75.

[8]We have employed a total of 48 NACE2 or NACE3 industry-specific price indices retrieved from the Eurostat New Cronos database, according to the classification reported in the Statistical Annex. The classification allows to divide industries into economies of scale, traditional, high tech and specialised industries, plus services. The same classification has been used by Davies and Lyons (1996) to divide industries into high, medium and low sunk costs. As such, the classification allows us to consider market structures, and hence prices, as relatively homogeneous within each industry.

[9]In a few cases (i.e. NACE16, 20 and 65) industries have displayed insufficient variation to identify the input coefficients. Accordingly, TFP measures from firms belonging to these industries have not been considered in the follow-up of our exercise.

unobservable productivity shocks.

Both methodologies have been employed in the literature, and both present some shortcomings. The LP methodology has been criticized on the grounds that the conditional demand for materials itself depends on the productivity shock, and thus materials are not a valid instrument to solve the simultaneity bias. The OP methodology, instead, does not suffer from the latter shortcoming, since the investment function is entirely determined before the productivity shock takes places; moreover, the OP approach offers a correction for the selection bias, incorporating in the algorithm a fitted value for the probability of exiting from the sample. However, a major assumption of the OP approach is the existence of a strictly monotonous relationship between the instrument (investment) and output. This means that any observation with zero or negative investment has to be dropped from the data. If the latter exclusion is significant, as it is typically the case in the early years of transition due to the substantial restructuring of the capital stock that has to be undertaken, the OP productivity estimates will be affected by an important selection bias. Since the latter is the case for a relevant share of our domestic firms, we have chosen to compute productivity through both approaches, in order to verify the extent to which the two methodologies yield different results for our purposes.

The analysis is presented in Table 2, which also includes the OLS estimates of TFP. Not surprisingly, the productivity levels estimated through OLS display the lowest values[10]. The point estimates of TFP calculated through the two semi-parametric methods also turn out to be different, thus showing the importance of the selection bias implicitly characterizing the OP approach in our sample[11]. However, as shown in Figure 1, it is important to notice that the distribution of domestic firms' TFP as retrieved through both the LP (unrestricted sample) and OP (restricted sample, positive investments) algorithms tend to overlap over the entire sampling period, once normalizing the TFP of a given firm by the industry average (correlation of 0.8, significantly different from zero at the 1 per cent level). Hence, any bias in the estimation of TFP eventually induced by either the LP or the OP methodology seems to cancel out, being it industry-specific and constant over time. Now, in a typical convergence regression the dependent variable (TFP) enters in the specification in first differences. Given the strong correlation between the two estimates, it follows that the choice of either methodology in the estimation of TFP is not likely to affect our results, as long as our dependent variable is considered in first differences.

[Table 2 and Figure 1 about here]

We have therefore opted for the LP procedure (see Annex 1 for further details) in order to derive TFP estimates for each firm, since the latter allows us to exploit all the data in our sample. Note also that we have run our estimates for domestic firms only, thus avoiding the possibility that the FDI status of a firm might have an effect on the choice of input factors,

[10]Typically, the simultaneity bias affecting OLS estimates leads to an upward bias in the estimate of the labor coefficient, which translates into a downward bias in the estimated TFP, since the latter is retrieved as the difference between the observed output and the predicted one.

[11]The OP algorithm has been calculated on the restricted sample of domestic firms displaying positive investments, while the LP algorithm has been calculated on the entire sample of firms.

another potential source of bias in the estimates of productivity (Van Biesebroeck, 2005).

Another important source of distortion in the estimation of TFP, not yet fully tackled by the convergence literature, relates to the so-called omitted price variable bias in the measurement of domestic firms' productivity. Since the seminal paper of Klette and Griliches (1996), it is known that proxying physical inputs and outputs through nominal variables deflated by a broad price index might lead to biased productivity measures, due to an omitted price variable bias induced by the correlation between (unobserved) individual firms' prices and their used inputs[12]. Such a bias can potentially affect the estimated TFP. The reason is that inputs are positively correlated with the level of output, which is typically negatively correlated with prices. If individual firm prices remain in the error term due to improper deflating, then the error term and the inputs are positively correlated, yielding an underestimated coefficient of labor and materials, and thus an overestimated TFP (thus opposite to the simultaneity bias one gets by using OLS, as previously discussed). As a result, controlling for the simultaneity bias, as it is nowadays common practice in the literature, but not for the omitted price variable bias, might lead to convergence estimates which are upward biased[13].

We assess these critiques in two ways: first of all, we follow Katayama, Lu and Tybout (2003), who argue that taking industry and region-specific averages on firm-specific TFP measures allows to partially counter the omitted price variable bias, since the cross-producer variation in productivity measures is much more problematic than the temporal variation of the population of plants. In addition, following the spirit of Klette and Griliches (1996), we control for the degree of imperfect competition on the demand side of the market allowing for spatial substitutability in demand (e.g. as in Syverson, 2005), assuming that deviations of domestic firms' prices of outputs and inputs (our measurement error) have a spatial component which can be controlled for. To this extent, we develop a slightly modified version of the original Levinsohn and Petrin (2003) algorithm, estimating an industry-specific production function augmented with regional fixed-effects, in order to pick up different pricing powers of domestic firms in the different Romanian regions (see Annex 1 for further details).

3.1 Convergence regressions

We start from a standard absolute convergence regression, where the change in the (log) TFP of a domestic firm i at time t is regressed against its level at time $t-1$ and a constant:

$$\Delta(\ln TFP_{it}) = \alpha + \beta \ln TPF_{it-1} + \varepsilon_{it} \quad (1)$$

To counter the omitted price variable bias, we have aggregated firm-specific TFP measures across NACE-3 industries and 8 regions over the years 1996-2001, as suggested by Katayama et

[12] Eslava et al. (2004) discuss this issue in their analysis of productivity of Colombian firms, where they can exploit the availability of firm-specific information on prices and quantities. DeLoecker (2005) provides a formal econometric discussion of the omitted price variable bias.

[13] Again, taking first differences of TFP as a dependent variable, the effect could disappear if the bias is constant over time, i.e. if individual firms price always at the same distance from the industry average price; the assumption is however very unlikely in the highly volatile context of transition.

al. (2003). To avoid possible problems induced by the non-normality of the TFP distribution, we have used as a dependent variable $\Delta \ln(\widetilde{TPF}_{zjt})$, i.e. the median, rather than the average, of $\Delta \ln TFP_{it}$ of each domestic firm i in industry z and region j, for a given year t:

$$\Delta(\ln \widetilde{TPF}_{zjt}) = \alpha + \beta \ln \widetilde{TPF}_{zjt-1} + \varepsilon_{zjt} \tag{2}$$

As we have argued, the latter treatment of the dependent variable yields a balanced panel across industries, regions and years, and allows us to minimize potential biases in our TFP measure deriving from the heterogeneity in the market power of individual firms.

Moving from absolute to conditional convergence, we have followed the related literature (Cameron et al., 2005) writing an equilibrium correction model (ECM) representation of a long-run cointegrating relationship between TFP in a non-frontier firm (or in a industry*region pair) and TFP in the frontier, for every given year. In this model, the TFP of every non-frontier firm (or industry*region pair) can grow either as a result of internal innovation or via technology transfers from the frontier, with the extent of the latter directly proportional to the 'distance' in terms of TFP separating the firm from the frontier. Among the possible covariates which might affect the rate of internal innovation of firms, we have included foreign direct investment (Griffith et al., 2006), the minimum efficient scale of each industry, the firms' absorptive capacity, and a variable measuring the average geographical distance between each region and the capital. The same variables are supposed to interact with the rate of technology transfers from the frontier. Finally, we have included in the cointegrating relationship a term in contemporaneous frontier growth, to control for possible technology shocks affecting the same frontier.

Our conditional convergence regression at the firm-level thus takes the form:

$$\begin{aligned}\Delta(\ln TFP_{izjt}) \quad = \quad & \alpha + \beta_1 \Delta \ln F_{zjt} + \beta_2 X_{zjt-1} + \beta_3 \ln\left(\frac{F_{zjt-1}}{TFP_{izjt-1}}\right) + \beta_4 X_{izjt-1} * \ln\left(\frac{F_{zjt-1}}{TFP_{izjt-1}}\right) \\ & + \gamma_{izj} + \gamma_t + \varepsilon_{zjt}\end{aligned} \tag{3}$$

where F_{zjt} is the TFP frontier, defined as the top 5% percentile of TFP of firms (both domestic and multinationals) in industry z and region j, for a given year t. Changes in the frontier are captured by the term $\Delta \ln F_{zjt}$. The variable $\left(\frac{F_{zjt-1}}{TFP_{izjt-1}}\right)$ measures the distance of each firm from the frontier, and it is supposed to capture technology transfers from the most productive firms. Our (lagged) covariates X_{zjt-1} responsible for internal TFP improvements by firms include Horizontal, Backward and Forward penetration indexes of MNEs, calculated from Input / Output tables as in Smarzynska Javorcik (2004); the minimum efficient scale (MES_{zt-1}) of industry z[14]; a proxy for domestic firms' absorptive capacity ($absorb_{zjt-1}$), measured as (the log of) domestic firms' average investment in intangible assets over total assets in a given industry/region at time $t-1$; and the (log of) geographic distance ($dist_z$) of each region from the capital city[15]. The same variables are interacted with the distance from the frontier, to

[14]The minimum efficient scale has been calculated as the median employment of the firms in each industry in a given year.

[15]The variable (taken in logs) is constructed taking the average of the distance (in km) between each county

check whether they affect the rate of technology transfers accruing to individual firms.

The aggregated version of the model used to counter the omitted price variable bias is then obtained by simply taking the median of equation (3) over the i firms.

4 Results

The results of the absolute convergence regressions (1) and (2) are presented in Table 3, based on a simple pooled OLS. We find evidence of absolute convergence across our population of firms, with the long run level of (log) TFP being equal to the value $-a/\beta$ in Equations (1) and (2). The value is similar when considering both multinational and domestic firms (left hand side of the Table) or domestic firms only. Moreover, we do not find a significant difference in terms of absolute convergence when considering average vs. median TFP values (Columns 1 vs. 2).

However, the long run TFP level is higher when firm-level measures of TFP are used (Columns 3), being almost twice the size of TFP taken as an average or median across sectors and regions (Columns 1 and 2). The latter finding is entirely consistent with our priors on the direction of the bias induced by the omitted price variable problem, as discussed in the previous Section: the firm-level TFP measure (1) actually picks up some effects induced by convergence in prices, rather than technology diffusion, with an upward bias resulting in the long-run convergence estimate of Equation (1).

Moving to the test for conditional convergence, Table 4 shows the results of the ECM (3) estimated using TFP changes of domestic firms as the dependent variable, and the presence of FDI as covariates, for each of our three measures of TFP: median across sectors and regions (Columns 1), firm-specific as from our LP estimates (Columns 2) and firm-specific retrieved from the modified LP algorithm discussed in the previous Section (Columns 3).

Domestic firms seem to benefit from technology transfers from the frontier, since the estimated coefficient on distance from the technological frontier is positive and highly statistically significant. Thus, consistent with the predictions of convergence theory, the further a firm lies behind the technological frontier, the higher its rate of TFP growth. The presence of FDI does not seem to have a robust effect neither on internal TFP improvements nor on the speed of convergence of domestic firms, once the distance from the frontier is taken into account (Columns 1a to 3a). Instead, FDI presence appears to be relevant when the latter term is omitted (Columns 1b to 3b). In other words, it seems that the traditional specification of the spillover literature (à la Smarzynska Javorcik, 2004) suffers from an omitted variable bias, that is distance from the technological frontier (which in our case includes both domestic and multinational firms). Moreover, only horizontal spillovers remain robust to our correction for the omitted price variable bias (Column 1b vs. Columns 2b and 3b), consistently with the fact that backward and forward linkages, being mediated by the market, are likely to be influenced by an imperfect deflationing of the production function.

The results also show that the estimated coefficient on contemporaneous frontier grow (the

town belonging to a given region and Bucuresti (the capital city).

term $\Delta \ln F_{zjt}$) is positive and statistically significant, a finding which, interpreted within our equilibrium correction model, signals the existence of a positive long-run cointegrating relationship between non-frontier and frontier TFP, in line with the previous result of absolute convergence[16].

Interestingly, also within the ECM specification the estimated rate of convergence is higher when measured using firm-level TFP (Column 2) vs. median TFP (Column 1), a result of the upward bias induced by the omitted price variable problem. Our correction for firm-level TFP retrieved using a modified LP algorithm (presented in Column 3) does not seem to solve the problem, yielding results very close to the standard LP estimation[17].

Another general problem possibly affecting the consistency of convergence regressions is given by the potential serial correlation in the error terms, which might bias the asymptotic properties of the estimators. Though the problem should be negligible in micro panels such as ours, characterized by a large number of cross-sectional units with respect to time, we have nevertheless performed a robustness checks of our specifications imposing an AR(1) structure in the error term, and reporting the Baltagi and Wu (1999) LBI test statistic in order to assess the extent of the problem. The results reported in Table 4 do not show significant problems of serial correlation[18].

Table 5 looks at the effects of other covariates on domestic firms TFP changes, always within our equilibrium correction model[19]. The overall results of positive technology transfers from the frontier and a positive long-run cointegrating relationship between non-frontier and frontier TFP remain valid, together with the finding that the estimated rate of convergence is higher when measured using firm-level TFP (Column 2) vs. median TFP (Column 1), due to the omitted price variable bias. We also find that a higher MES tends to stimulate internal gains in TFP by firms, as firms in industries where firms are on average larger are more likely to possess a sufficient level of absorptive capacity (Aitken and Harrison, 1999). When absorptive capacity is measured as the ratio of intangible to tangible assets, instead, the effect is not significant. In terms of rate of technology transfer, the latter tend to decrease in line with the the average distance of each firm from the capital region[20]. Finally, also in Table 5 our correction for firm-level TFP retrieved using a modified LP algorithm (presented in Column 3) yields results very close to the standard LP estimation, while the Baltagi and Wu (1999) LBI test statistic does not reveal major problems of serial correlation.

In Table 6 we analyze the second moment of our TFP distribution, running a σ-convergence regression where the standard deviation of firm-specific (log) TFP has been calculated for each

[16] For the exact relationship between the estimated ECM and the concepts of β and σ-convergence, see Cameron et al. (2005).

[17] Note that our correction of the LP algorithm assumes that deviations of domestic firms' prices from the industry average (our measurement error generating the omitted price variable bias) have only a spatial component, which can thus be controlled for through a proper set of region-specific fixed effects. It ignores instead other possible sources (e.g. industry-specific) of market power.

[18] A value of the statistic around 2 signals the absence of serial correlation in the residuals.

[19] We discuss here only the covariates whose results are robust across the three different specifications.

[20] The result is in line with the spatial distribution of regional income in Romania, where the capital region clearly outweighs the other in terms of per capita GDP.

industry, region and year[21]. Our dependent variable $\sigma(TFP_{zjt})$ has been regressed against a time trend, industry and region fixed-effects, and the interaction of the trend with our covariates (MES, absorption capacity and distance to the capital), in order to assess their impact on this alternative measure of convergence. The results, reported in Table 6, do not find evidence of σ-convergence or divergence when all firms are considered (Column 1), even when controlling for sector and region fixed effects (Column 2). In the latter specification, absorptive capacity measures as the ratio of intangibles to tangible assets makes divergence more likely, in line with the results of Basu and Weil (1998), which use the concept of 'appropriate technology' as an explanation for divergence in growth. Finally, we do find evidence of (weak) σ-divergence when considering only domestic firms.

5 Conclusions and policy implications

In this paper we take a microfounded perspective to test new explanations for the speed of convergence of firms' productivity. Firm, sectoral and regional data available over a time span of seven years in Romania allow in fact for a more direct test of the significance of potentially important variables on the speed of convergence with respect to country specific data. At the same time, firm-specific measures of TFP allow us to correct for a number of biases (simultaneity and omitted price variable) potentially affecting the convergence measure.

We show that, in line with our theoretical priors, failing to take into account the omitted price variable bias in TFP measures results in a significant (more than 60%) upward bias in the detected speed of absolute convergence. In terms of conditional convergence, we employ an equilibrium correction model, as in Cameron et al. (2005), to show that domestic firms seem to benefit from technology transfers from the frontier, with the firms lying further from the technological frontier having a higher rate of TFP growth, in line with convergence theory. In terms of other covariates affecting convergence, the presence of FDI in horizontal, backward or forward industries does not seem to have a robust effect on the TFP convergence of domestic firms once distance from the frontier is taken into account, a result which sheds some new light on the spillover literature (e.g. Smarzynska Javorcik, 2004). We also find that a higher MES tends to stimulate internal gains in TFP by firms, as argued by Aitken and Harrison (1999). In terms of rate of technology transfers from the frontier, the latter tend to decrease in line with the the average distance of each firm from the capital region, characterized by a significantly higher level of per capita GDP. Finally, we do not find evidence of σ-convergence or divergence when all firms are included in the analysis, with only a slight σ-divergence detected for the TFP of domestic firms.

POLICY IMPLICATIONS [...]

[21] It is well known in the literature that β-convergence is a necessary, but not sufficient condition for σ-convergence.

References

Aghion, P., P. Howitt and D. Mayer-Foulkes (2005), The Effect of Financial Development on Convergence: Theory and Evidence, Quarterly Journal of Economics 120, 173-222.

Aitken, B.J. and A.E. Harrison (1999), Do Domestic Firms Benefit from Direct Foreign Investment? Evidence from Venezuela, American Economic Review 89, 605-618,

Baltagi, B.H. and P. Wu (1999), Unequally Spaced Panel Data Regression with AR(1) Disturbances, Econometric Theory 15, 814-823.

Basu, S. and D. N. Weil (1998), Appropriate Technology And Growth, The Quarterly Journal of Economics 113: 1025-1054.

Benhabib, J. and M. Spiegel (2005), Human Capital and Technology Diffusion, in: Handbook of Economic Growth, Vol. 1, part 1, P. Aghion and S. Durlauf (Eds.), Elsevier, 935-966, Amsterdam.

Bernard, A.B. and C.I. Jones (1996), Comparing Apples to Oranges: Productivity Convergence and Measurement Across Industries and Countries, American Economic Review 86: 1216-1238

Cameron, G., J. Proudman and S. Redding (2005), Technological Convergence, R&D, Trade and Productivity Growth, European Economic Review 49, 775-807.

Griffith, R., S. Redding and J. Van Reenen (2004), Mapping the Two Faces of R&D: Productivity Growth in a Panel of OECD Countries, Review of Economics and Statistics 86, 883-895.

Griffith, R., S. Redding and H. Simpson (2006), Technological Catch-up and the Role of Multinationals, revised version of CEPR Discussion Paper No. 3765, July 2006.

Griliches, Z. and J. Mairesse (1995), Production Functions: the Search for Identification, NBER Working Paper No. 5067, Washington, D.C.

Islam, N. (1995), Growth Empirics: A Panel Data Approach, Quarterly Journal of Economics 110, 1127-1170.

Katayama, H. S. Lu and J.R. Tybout (2003), Why Plant-Level Productivity Studies Are Often Misleading, and an Alternative Approach to Inference, NBER Working Paper No. 9617, Washington D.C.

Klette, T.J. and Z. Griliches (1996), The Inconsistency of Common Scale Estimators When Output Prices Are Unobserved and Endogenous, Journal of Applied Econometrics 11, 343-361.

Lee, K., M.H. Pesaran and R. Smith (1998), Growth Empirics: A Panel Data Approach – A Comment, Quarterly Journal of Economics 113, 319-323.

Levinsohn, J. and A. Petrin (2003), Estimating Production Functions Using Inputs to Control for Unobservables, Review of Economic Studies 70, 317-342.

Mohnen, P. (1996), R&D Externalities and Productivity Growth, STI Review 18, 39-66.

Olley, S. and A. Pakes (1996), The dynamics of productivity in the telecommunications equipment industry, Econometrica 64, 1263-98.

Temple, J. (1999), The New Growth Evidence, Journal of Economic Literature 37, 112-156

Smarzynska Javorcik, B. (2004), Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers through Backward Linkages, American Economic Review 94, 605-627.

Syverson, C. (2005), Market Structure and Productivity: A Concrete Example, Journal of Political Economy 112, 1181-1222.

Van Biesebroeck, J. (2005), Exporting Raises Productivity in sub-Saharan African Manufacturing Firms, Journal of International Economics 67, pp. 373-391

Table 1. The census of Romanian firms in Manufacturing and Construction (1996-2001, number of firms and rates)

Year	Sample Stock (AMADEUS)	Official Stock	Sample Coverage
1996	39799	41228	0.97
1997	43593	45432	0.96
1998	47491	49324	0.96
1999	50257	52295	0.96
2000	50246	53568	0.94
2001	48718	57086	0.85

of which:

	Domestic firms			Multinational firms					
Year	Entry	Exit	Active Firms	Entry	Exit	Active Firms	MNEs Penetration	Entry Rate	Exit Rate
1996			36634			3165	0.08		
1997	4771	1576	39829	728	129	3764	0.09	0.14	0.04
1998	5006	1827	43008	880	161	4483	0.09	0.14	0.05
1999	4606	2685	44929	1048	203	5328	0.11	0.12	0.06
2000	2514	3422	44021	1212	315	6225	0.12	0.07	0.07
2001	2228	4268	41981	1234	722	6737	0.14	0.07	0.10

Percentage of industry distribution over total sample:

	1996			2001		
NACE2	**All Firms**	**Dom**	**MNEs**	**All Firms**	**Dom**	**MNEs**
15	25.5%	25.4%	27.7%	22.5%	22.9%	19.8%
17	4.4%	4.4%	4.4%	3.9%	3.8%	5.1%
18	8.0%	8.2%	6.5%	7.7%	7.5%	9.4%
19	2.3%	2.2%	3.8%	2.6%	2.1%	5.6%
20	7.9%	7.9%	7.6%	8.4%	8.1%	10.4%
21	1.0%	0.9%	1.9%	1.0%	0.9%	1.7%
22	5.2%	5.1%	6.5%	5.4%	5.5%	4.7%
24	2.0%	1.9%	3.5%	2.1%	1.9%	3.1%
25	3.1%	2.9%	4.4%	3.0%	2.7%	4.5%
26	2.6%	2.6%	2.8%	2.7%	2.7%	3.1%
27	0.7%	0.7%	1.2%	0.8%	0.7%	1.2%
28	5.7%	5.9%	4.5%	6.0%	6.1%	5.3%
29	1.5%	1.4%	3.0%	1.7%	1.5%	3.1%
30	0.8%	0.7%	2.1%	0.9%	0.8%	1.2%
31	1.1%	1.1%	1.7%	1.2%	1.0%	1.8%
32	0.3%	0.3%	0.9%	0.3%	0.3%	0.7%
33	1.0%	1.0%	1.3%	1.0%	1.1%	0.9%
34	0.5%	0.5%	0.9%	0.6%	0.5%	0.9%
35	0.4%	0.4%	0.5%	0.5%	0.4%	0.7%
36	5.1%	5.1%	5.2%	5.3%	5.2%	5.8%
45	20.7%	21.7%	9.7%	22.3%	24.1%	11.0%
Total firms	**39799**	**36634**	**3165**	**48718**	**41981**	**6737**

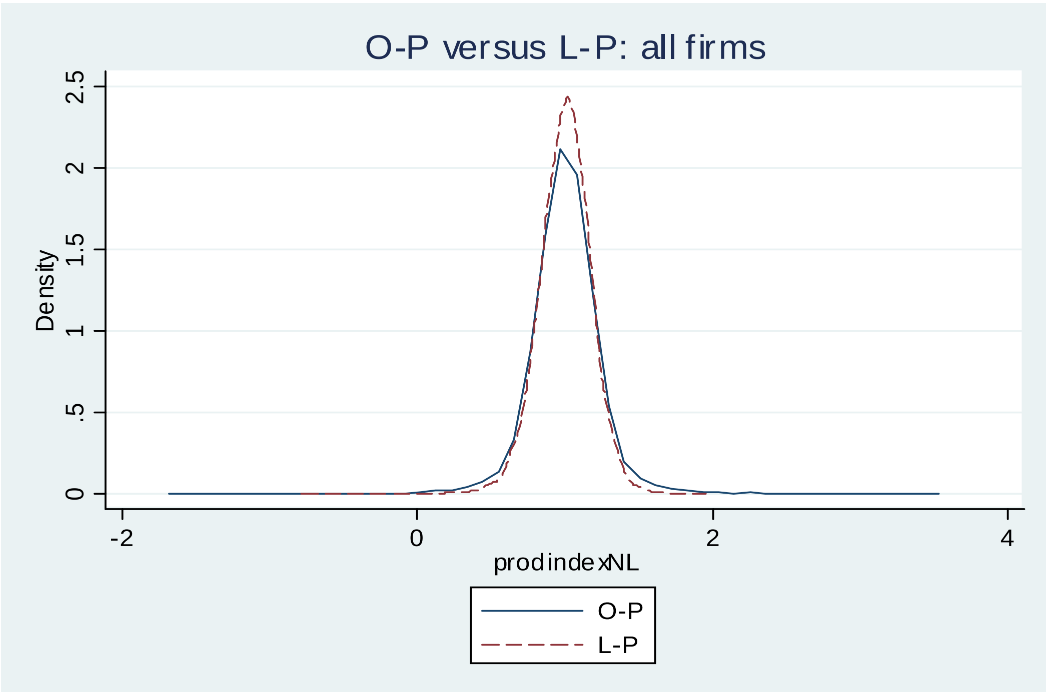
Source: author's elaboration from Amadeus data

Table 2. TFP measures

NACE Industry	TFP OLS	Std. Dev.	TFP L-P	Std. Dev.	TFP O-P	Std. Dev.	N. of obs. (unrestricted)
158	1.939	0.911	5.216	1.155	1.630	0.913	3337
22	2.116	0.948	3.317	1.025	5.692	1.435	1913
156	2.152	0.915	7.015	1.328	6.119	1.181	716
45	2.385	0.850	5.117	1.053	3.124	0.865	8886
153, 155	2.417	0.901	4.620	1.065	2.289	0.905	848
361, 362	2.442	0.847	5.177	1.067	6.354	1.207	1991
151, 152	2.452	0.927	5.083	1.130	2.761	0.930	1413
159	2.456	1.052	3.481	1.123	4.306	1.166	637
243, 245	2.687	1.017	4.823	1.186	5.196	1.200	608
28	2.816	0.928	4.071	0.986	3.858	0.963	2883
26	2.877	0.882	6.484	1.215	7.691	1.404	960
252, 262	2.880	0.890	5.066	1.062	5.769	1.137	1091
17	3.142	0.880	5.131	1.020	8.453	1.643	1645

Note: Log TFP for each industry calculated as a weighted average of individual domestic firms' TFP, estimated through OLS, Levinsohn-Petrin (2003) or Olley-Pakes (1996) semiparametric algorithm for the years 1995-2001. In the O-P algorithm, only the sample of firms displaying non-zero investment has been considered (42% of available firms, on average).

Figure 1. Distribution of estimated domestic firms' productivity: O-P vs. L-P



Note: TFP index of individual domestic firms for the period 1995-2001, normalized to industry average in a given year. Olley and Pakes, 1996 (O-P) estimates are performed on the restricted sample (only domestic firms displaying positive investments), while the Levinsohn and Petrin, 2003 (LP) estimates are performed over the entire sample for the corresponding industries.

Table 3. Absolute and Conditional Convergence

Dep var: $\Delta(\ln TFP_{izjt})$	All firms (1)	All firms (2)	All firms (3)	Domestic firms (1)	Domestic firms (2)	Domestic firms (3)
$\ln TFP_{izjt-1}$	-.12***	-.11***	-.18***	-.12***	-.11***	-.18***
	(.005)	(.005)	(.002)	(.005)	(.005)	(.002)
constant	.03***	.02***	.05***	.02***	.02***	.05***
	(.004)	(.004)	(.001)	(.004)	(.004)	(.001)
FE	no	no	no	no	no	no
R^2	.12	.10	.12	.12	.11	.12
N. of obs.	4109	4109	98880	3930	3930	85820

***, ** significant at the 1 or 5 per cent level, respectively.

(1) Average TFP of region / industry in a given year.

(2) Median TFP of region / industry in a given year.

(3) Firm-specific TFP.

(4) Median TFP.of region / industry in a given year with industry* region fixed-effects.

Table 4. ECM of domestic TFP growth and FDI presence

Dep var: $\Delta(\ln TFP_{izjt})$	(1a)	(2a)	(3a)	(1b)	(2b)	(3b)
$\Delta \ln F_{zjt}$	.23***	.42***	.41***			
	(.009)	(.006)	(.006)			
Horizontal FDI $_{zjt-1}$	-.04	-.004	-.01	.07***	.03***	.03***
	(.064)	(.022)	(.021)	(.021)	(.006)	(.006)
Backward linkages $_{zjt-1}$	-.007	.06	.08**	.005	-.04***	-.04***
	(.115)	(.041)	(.041)	(.031)	(.010)	(.010)
Forward linkages $_{zjt-1}$	-.10	-.03	-.02	-.11*	-.04***	-.04***
	(.209)	(.075)	(.074)	(.062)	(0.20)	(0.20)
$\ln(F_{zjt-1}$ / $TFP_{izjt-1})$	.21***	.73***	.73***			
	(.028)	(.014)	.(014)			
Hor$*\ln(F_{zjt-1}$ / $TFP_{izjt-1})$	.03	-.05**	-.04*			
	(.058)	(.022)	(.023)			
Bwd$*\ln(F_{zjt-1}$ / $TFP_{izjt-1})$	.06	-.14***	-.19***			
	(.094)	(.040)	(.041)			
Fwd$*\ln(F_{zjt-1}$ / $TFP_{izjt-1})$	-.03	.16*	.17**			
	(.188)	(.080)	(.081)			
industry * region FE	yes	yes	yes	yes	yes	yes
time FE	yes	yes	yes	yes	yes	yes
constant	-.17***	-.49***	-.49***	-.07***	-.06***	-.06***
	(.035)	(.012)	(.012)	(.014)	(.005)	(.005)
R^2	.17	.13	.13	.05	.01	.01
Baltagi-WU LBI	2.36	2.30	2.30			
N. of obs.	3517	61986	61986	3517	61986	

***, ** significant at the 1 or 5 per cent level, respectively. FE within estimator.

Standard errors clustered on individual observational units.

(1) Median TFP of region / industry in a given year.

(2) Firm-specific TFP. Standard errors clustered on individual observational units.

(3) Firm-specific TFP retrieved from a modified version of the original Levinsohn and Petrin (2003) algorithm, estimating an industry-specific production function augmented with regional fixed-effects.

Table 5. ECM of domestic TFP growth and firms' characteristics

Dep var: $\Delta(\ln TFP_{izjt})$	(1)	(2)	(3)
$\Delta \ln F_{zjt}$	.24***	.44***	.43***
	(.009)	(.006)	(.006)
MES $_{zjt-1}$	.01***	.04***	.03***
	(.005)	(.005)	(.005)
Absorptive capacity $_{zjt-1}$	-.34	-.08	-.11
	(.359)	(.148)	(.147)
$\ln(F_{zjt-1}\ /\ TFP_{izjt-1})$	.32***	.84***	.83***
	(.029)	(.019)	(.019)
Dist$*\ln(F_{zjt-1}\ /\ TFP_{izjt-1})$	-.014***	-.008***	-.008***
	(.005)	(.003)	(.003)
MES$*\ln(F_{zjt-1}\ /\ TFP_{izjt-1})$	-.01*	-.03***	-.03***
	(.007)	(.007)	(.007)
Absorb$*\ln(F_{zjt-1}\ /\ TFP_{izjt-1})$	-2.20**	.01	.02
	(.962)	(.055)	(.051)
industry * region FE	yes	yes	yes
time FE	yes	yes	yes
constant	-.19***	-.55***	-.55***
	(.013)	(.012)	(.012)
R^2	.18	.14	.14
Baltagi-WU LBI	2.34	2.28	2.29
N. of obs.	3930	85820	85820

***, ** significant at the 1 or 5 per cent level, respectively. FE within estimator.

Standard errors clustered on individual observational units.

(1) Median TFP of region / industry in a given year.

(2) Firm-specific TFP. Standard errors clustered on individual observational units.

(3) Firm-specific TFP retrieved from a modified version of the original Levinsohn and Petrin (2003) algorithm, estimating an industry-specific production function augmented with regional fixed-effects.

Table 6. Testing for σ-convergence

Dep var: $\sigma(\ln TFP_{zjt})$	(1)	(2)	(3)	(4)
trend$_t$	.001	.003	.003***	.008***
	(.001)	(.004)	(.001)	(.004)
trend$_t * MES_{zt-1}$		-.001		-.002**
		(.001)		(.001)
trend$_t * absorb_{zjt-1}$		.056**		-.046
		(.028)		(.032)
trend$_t$*ln$dist_j$		.001		.001
		(.001)		(.001)
_cons	.39***	.39***	.35***	.34***
	(.004)	(.005)	(.004)	(.005)
industry * region FE	no	yes	no	yes
R^2	.01	.01	.01	.01
N. of obs.	4378	3789	4131	3572

***, ** or * significant at the 1, 5 or 10 per cent level, respectively.

(1) & (2) Standard deviation of TFP calculated for both domestic firms and MNEs.

(3) & (4) Standard deviation of TFP calculated for domestic firms only.

Annex 1: Levinsohn and Petrin (2003) productivity estimates

Let y_t denote (the log of) a firm's output in a Cobb-Douglas production function of the form

$$y_t = \beta_0 + \beta_l l_t + \beta_k k_t + \beta_m m_t + \omega_t + \eta_t \tag{A1.1}$$

where l_t and m_t denote the (freely available) labour and intermediates inputs in logs, respectively, and k_t is the logarithm of the state variable capital. The error term has two components: η_t, which is uncorrelated with input choices, and ω_t, a productivity shock unobserved to the econometrician, but observed by the firm. Since the firm adapts its input choice as soon as she observes ω_t, inputs turn out to be correlated with the error term of the regression, and thus OLS estimates of production functions yield inconsistent results. To correct for this problem, Levinsohn and Petrin (2003), from now on LP, assume the demand for intermediate inputs m_t (e.g. material costs) to depend on the firm's capital k_t and productivity ω_t, and show that the same demand is monotonically increasing in ω_t. Thus, it is possible for them to write ω_t as $\omega_t = \omega_t(k_t, m_t)$, expressing the unobserved productivity shock ω_t as a function of two observables, k_t and m_t. To allow for identification of ω_t, LP follow Olley and Pakes (1996) and assume ω_t to follow a Markov process of the form $\omega_t = E[\omega_t|\omega_{t-1}] + \xi_t$, where ξ_t is a change in productivity uncorrelated with k_t. Through these assumptions it is then possible to rewrite Equation (A1.1) as

$$y_t = \beta_l l_t + \phi_t(k_t, m_t) + \eta_t \tag{A1.2}$$

where $\phi_t(k_t, m_t) = \beta_0 + \beta_k k_t + \beta_m m_t + \omega_t(k_t, m_t)$. By substituting a third-order polynomial approximation in k_t and m_t in place of $\phi_t(k_t, m_t)$, LP show that it is possible to consistently estimate the parameter $\widehat{\beta}_l$ and $\widehat{\phi}_t$ in Equation A1.2. For any candidate value β_k^* and β_m^* one can then compute a prediction for ω_t for all periods t, since $\widehat{\omega}_t = \widehat{\phi}_t - \beta_k^* k_t - \beta_m^* m_t$ and hence, using these predicted values, estimate $E[\omega_t|\widehat{\omega_{t-1}}]$. It then follows that the residual generated by β_k^* and β_m^* with respect to y_t can be written as

$$\widehat{\eta_t + \xi_t} = y_t - \widehat{\beta}_l l_t - \beta_k^* k_t - \beta_m^* m_t - E[\omega_t|\widehat{\omega_{t-1}}] \tag{A1.3}$$

Equation (A1.3) can then be used to identify β_k^* and β_m^* using the following two instruments: if the capital stock k_t is determined by the previous period's investment decisions, it then does not respond to shocks to productivity at time t, and hence $E[\eta_t + \xi_t|k_t] = 0$; also, if the last period's level of intermediate inputs m_t is uncorrelated with the error period at time t (which is plausible, e.g. in the case of material costs), then $E[\eta_t + \xi_t|m_{t-1}] = 0$. Through these two moment conditions, it is then possible to write a consistent and unbiased estimator for β_k^* and β_m^* simply by solving

$$\min_{(\beta_k^*, \beta_m^*)} \sum_h [\sum_t (\widehat{\eta_t + \xi_t}) Z_{ht}]^2 \tag{A1.4}$$

with $Z_t \equiv (k_t, m_{t-1})$ and h indexing the elements of Z_t.

With specific reference to our exercise, note however that the intercept β_0 of the production function is not separately identified in the estimation. In our robustness checks, we have thus modified the procedure described above by incorporating in Equation (A1.2) regional fixed-effects. As a result, we can retrieve firm-specific TFP measures corrected for region-specific factors which might affect the pricing power of domestic firms.