Skill Biased Technical Change and Labor Market Inefficiency*

April 6, 2022

Michele Battisti
University of Palermo, RCEA (michele.battisti@unipa.it)

Massimo Del Gatto
"G.d'Annunzio" University, LUISS Guido Carli, CRENoS (massimo.delgatto@unich.it)

Christopher F. Parmeter
University of Miami (cparmeter@bus.miami.edu)

Abstract

Economic literature has so far produced very limited (country-level only) evidence on the magnitude of skilled biased technical change (SBTC) and not investigated at all the extent to which, coupled with labor market imperfections, it can be associated to an inefficient use of labor. We present a novel approach to estimating SBTC from the production side (that is, independently from observed wages), thereby allowing for the presence of, and the evaluation of, labor market inefficiency. Using WIOD (World Input-Output Database) data (38 countries, 30 sectors), this approach allows us to provide the first country-sector evidence on SBTC over the 1995-2005 decade and to quantify the extent of economic inefficiency, arising from labor market imperfection, in terms of discrepancy between marginal rate of technical substitution (i.e. relative productivity) and relative wage of skilled and unskilled workers. On average, we find the productivity of skilled workers to grow by around 11.5% more than that of the unskilled, mostly driven by SBTC (13.7%), rather than factor accumulation (-2.1%). Economic inefficiency decreases by 10.5% on average and labor market rigidities are more binding in sectors that are more exposed to technological change. Substantial heterogeneity emerges. While displaying the third highest SBTC figure (29%), USA features a relative decrease in the productivity of skilled workers.

Keywords: Skill Biased Technical Change, Technological Progress, Production Function Estimation, Wage Premium.

JEL Classification: D24, E24, J24, O33.

*We thank Francesco Caselli, Chris Papageorgiou, David Rivers, Federico Rossi and Joseph Zeira for helpful comments. We are grateful to participants at the 2018 North American Productivity Workshop in Miami, the 2019 RIDGE Workshop on International Trade and Development in Macroeconomics in Montevideo, the 2020 CEPR Structural Transformation and Economic Growth Workshop, the 2020 ESCoE Conference on Economic Measurement and the 2020 CEPR Macroeconomics and Growth Meeting. We are also indebted to a number of academic seminar participants for insightful discussions.

1 Introduction

The economic implications of technological progress are far from **unequivocal assessment**. An issue that still deserves attention, **notably at the policy level (e.g., OECD, 2018),**[1] is the extent to which technological advancement in recent decades happened in a "skill biased" form. When technical change comes with increasing marginal productivity (MP - we hereinafter use MP as an abbreviation for marginal product/productivity) of skilled workers, relative to unskilled workers (i.e., "skill biased technical change" - SBTC),[2] its asymmetric effects on the relative demand of skilled and unskilled labor are likely to transmit to relative wages (i.e., skill, or wage, premium) and, finally, to economic efficiency.[3] In fact, by increasing the MP of skilled labor relative to the MP of unskilled labor, SBTC requires adjustment in terms of either the wage ratio or the relative use of skilled and unskilled labor (i.e., the labor ratio), or both. As long as labor market frictions/imperfections affect the skilled and unskilled labor markets asymmetrically and prevent the adjustment from happening, technological progress can be detrimental to economic efficiency, and crucially affect wage distribution. This can dampen, even substantially, the output gains associated with technological progress and it might partially explain the missing burst in productivity and output corresponding to "waves" of technological progress, such as the "computer era" in the 70s and the 80s (Solow, 1987), and the advent of artificial intelligence in the last decade (Brynjolfsson, 1993; Brynjolfsson et al., 2019, 2021).

Indeed, apart from a few contributions (i.e., Caselli and Coleman, 2006; Caselli, 2016) showing SBTC to be larger in higher-income countries than in lower-income countries (with the US being the leader, in both the 1990s and 2000s), the empirical evidence on SBTC is scant. In particular, there is no evidence at all on the magnitude of the SBTC at a country-sector level and, as a consequence, there is lack of consensus about the skill biased nature of the above "waves" of technological progress. While several studies (Bound and Johnson, 1992; Katz and Murphy, 1992; Levy and Murname, 1992; Juhn et al., 1992) established a wide consensus that US wage ratios had been growing sharply in the 1980s and that the leading factor of growth was the increase in the relative demand for skill (Lemieux, 2008), "*the apparent stability of aggregate wage inequality over the 1990s presents a potentially important puzzle for the SBTC hypothesis, since there were continuing advances in computer-related technology throughout the decade that were arguably as skill biased as the innovations in the early 1980s*" (Card and Di Nardo, 2002, p. 748). Moreover, the idea that technological change was the major determinant of the wage inequality pattern in the US since 1970 (see Zeira, 1998 and Acemoglu, 2002) clashes with the evidence that

[1] OECD (2018) documents a 7% overall increase in measured decoupling between labor productivity and real wages, which is equally attributed to increasing wage inequality and labor accumulation.

[2] An alternative explanation in terms of routine biased technical change (RBTC) was proposed in Autor et al., 2003 and investigated in Goos and Manning, 2007, Goos et al., 2014 and others.

[3] In perfectly competitive markets, firms choose the amount of each input to deploy by equalizing its marginal revenue product to its price **(i.e., "economic efficiency" condition)**. In relative terms, this entails that the marginal rate of technical substitution (MRTS) between any two inputs (the ratio between the two marginal revenue products) equals the corresponding price ratio. When technological progress implies a uniform "shift" of the production function, the MP of all inputs is affected in the same proportion ("Hicks-neutral" technological progress). In this case, the economic efficiency condition is unaffected and no adjustment in the relative demand of inputs occurs. Differently, when technological progress determines a change in the "curvature" of the production function, the MP of inputs is affected in different proportions (i.e., "factor biased" technical change), and the MRTS changes at given input levels. In this case, the relative demand of inputs is crucially affected, and with it, the input price ratio, as long as the change in relative demand is not compensated by a proportional change in input relative supply.

industrialized economies such as France, Germany, Italy and Japan did not experience any significant growth in wage inequality, despite their exposure to similar technology shocks as the US.

In this paper, we argue that such a puzzle exists because available SBTC studies do not address SBTC from the "production side" but, instead, from the "wage side". Indeed, SBTC estimation would require estimating the MP of skilled and unskilled labor from the production function, computing the MRTS as MP ratio and, finally, isolating the MRTS variation not associated to changing the used amount of the two types of labor, as well as of capital. However, estimating a country by sector by year specific production function using country by sector by year data is a daunting task, as the number of production function parameters to be estimated equals the number of observations. Put it simply, this is the reason why available SBTC literature follows an "indirect approach", in which SBTC is residually obtained "from" observed skill premia (hence, "wage-based" approach), once the relative skill supply of workers is accounted for, under the hypothesis of perfectly competitive labor markets.[4]

In providing with the only available evidence on SBTC at the country level, Caselli (2016) points out how important is "for this methodology that relative wages are informative about relative marginal productivities". For example, "if social and political pressures for containing wage dispersion are much more severe in rich than in poor countries [...] this type of measurement error will bias the results against a finding of skill bias" (Caselli, 2016, section 2.1). Hence, by retrieving SBTC from the economic efficiency condition arising under the hypothesis of perfectly competitive labor market, available SBTC studies tend to conflate "true" SBTC and labor market inefficiencies associated with the action of frictions preventing relative wages from adjusting to the variation in the relative supply and demand of skills (Freeman and Katz, 1995), and vice-versa. The recent literature on misallocation has largely documented the presence of deviations from the efficiency condition (Restuccia and Rogerson 2008, 2013; Hsieh and Klenow, 2009; Bartelsman, et al., 2013; Calligaris et al., 2018) driven by the action of rigidities in the capital and labor markets. Labor market frictions can arise at different levels. Among others, Card (1992) and Freeman (1993) focus on unionization, while Di Nardo et al. (1996) and Lee (1999) highlight the role of minimum wage.

While such a mis-measurement is to some extent inescapable (Caselli, 2016) in a standard parametric context,[5] it actually hampers the scope for drawing well targeted policy conclusions as long as it rules out the chance to study the extent to which labor market imperfections can eventually prevent skill premia and/or firms' hired quantities of skilled and unskilled workers from freely adjusting to a mutant technological environment, thereby driving even large deviations from the economic efficiency condition.

To overcome these limitations, we take advantage of recent developments in nonparametric estimation (see Li and Racine, 2007 and Henderson and Parmeter, 2015 for textbook treatments) to estimate the MP of inputs from aggregate international data including country, sector and time effects. This allows

[4]With perfectly competitive labor markets, the MRTS between skilled and unskilled labor equals the corresponding wage ratio. Hence, the change in the MRTS can be inferred from the observed change in the skill premium. Since the change in MRTS only depends on SBTC, and the change in the labor ratio (i.e., ratio of deployed quantity of skilled labor to deployed quantity of unskilled labor), the SBTC can be "residually" retrieved from the (observable) change in the discrepancy between the wage ratio and the labor ratio, once the skilled and unskilled labor aggregates are properly constructed and the skill premium is estimated (see Caselli, 2016, on these two key steps).

[5]See Dorazelski and Jaumandreu (2018) for a semi-parametric multi-step approach at the firm level.

us to estimate the MRTS between any two inputs (skilled and unskilled labor in the application) directly from the production function and isolate a SBTC component, net of factor accumulation (FA) effects, through counterfactual analysis. As a result, we are able to quantify the wedge between MRTS and wage ratio and to understand whether its eventual increase is determined by SBTC coupled with stagnating wage ratios, a circumstance that can arguably be attributed to the presence of frictions affecting the skilled and unskilled labor markets asymmetrically.

Using country-sector information drawn from the WIOD database (40 countries at the 2-digit sectoral level), we find that the MRTS between skilled and unskilled labor has been growing at a yearly rate of 1.15% on average, over the 1995-2005 decade. Overall, most of this change is driven by SBTC, for which we report an average yearly contribution of 1.37%, and much less by FA, for which we estimate a slightly negative contribution (-0.21%) to the MRTS variation. Average values mask substantial heterogeneity across countries and sectors, with the FA effect sometimes dominating the SBTC component. Notably, this is the case in the US, for which we estimate the second last MRTS growth in the decade (-19.5%) associated with a conspicuously negative FA contribution (around -49%) and a positive SBTC (+29%), the third highest figure. In a similar situation of negative MRTS variation we find developed countries like Korea, Japan and, to a lesser extent, Germany and Great Britain. On the opposite side, it is possible to identify a group of less developed countries, whose relatively high MRTS increase is mainly driven by the FA component. Notably, this is the case of India, the third country in terms of MRTS growth, featuring an even negative SBTC.

We then consider the change in the discrepancy between the MRTS and the wage ratio of high to low skilled labor. For this measure of economic (in)efficiency, we document an overall 10.5% decrease. Although with considerable cross-country and cross-sector variability, we find the skill premium to grow less than the MRTS in the majority of countries (notably, Korea, the US, Japan Germany are among the few countries in which the skill premium has grown more than the MRTS) and sectors.

We report econometric analysis suggesting that economic efficiency can be bigger in presence of labor market institutions that reduce wage differentials.

We finally use counterfactual analysis to show that full adjustment of wages to MP would have increased within-country wage inequality by 10 points in terms of the Gini index, mostly through further decreases in low wages.

A policy message emerges insofar the analysis shows that skilled workers are not fully reaping the benefits of technological progress, compared to unskilled workers.

The exposition is organized as follows. Section 2 gives a simplified presentation of SBTC, while Section 3 discusses the indirect approach. Section 4 presents our direct approach. Section 5 reports the SBTC analysis. Section 6 carries out the economic efficiency analysis and presents counterfactual analysis on wage adjustment. Section 7 concludes. Finally, Appendix A explains the empirical methodology in more details, Appendix B discusses endogeneity, Appendix C presents the overall distributions featuring our estimated MPs.

2 Definition of terms: SBTC and economic inefficiency

Let us begin with a general production function (in logarithmic form) of country c, at time t (for brief notation, we omit the industry index, considering that everything is also specific to a given sector)

$$y_{cj,t} = m_{cj,t}(\boldsymbol{x}_{cj,t}) + z_{cj,t}. \tag{1}$$

Here, country c's log-output in sector j, at time t, $y_{cj,t}$, depends on the contemporaneous log-values of capital $k_{cj,t}$, skilled labor $s_{cj,t}$ and unskilled labor $u_{cj,t}$, through production technology $m_{cj,t}(\cdot)$, as well as on $z_{cj,t}$, which captures Total Factor Productivity (TFP) and idiosyncratic productivity shocks.

According to (1), cross-country differences in the production technology are captured by $m_{cj,t}$, with $z_{cj,t}$ accounting for cross-country and cross-sector output differences not explained by different input choices ($k_{cj,t}$, $s_{cj,t}$, $u_{cj,t}$) and/or production technology. Given the amount of inputs, country c's actual technology at time t can make it more productive than a generic country f at time t (i.e., $m_{cj,t} > m_{fj,t}$) or more productive than itself at $t-1$ (i.e., $m_{cj,t} > m_{cj,t-1}$).

In this framework, the MRTS can be defined as

$$MRTS^{s,u}_{cj,t} = \frac{m^s_{cj,t}(\boldsymbol{x}_{cj,t})}{m^u_{cj,t}(\boldsymbol{x}_{cj,t})} \tag{2}$$

with $m^s_{cj,t}$ and $m^u_{cj,t}$ denoting the MP of skilled and unskilled labor, respectively.

Hence, MRTS varies because of both technology (i.e., SBTC) and amount of factors (i.e., factor accumulation). To disentangle these two components (see Section 4.2), let us use

$$\Delta\mathbb{B}^{s,u}_{cj,tT} = \left(MRTS^{s,u}_{cj,T} - MRTS^{s,u}_{cj,t}\right) \Big| \, (k_{cj,T} = k_{cj,t}, s_{cj,T} = s_{cj,t}, u_{cj,T} = u_{cj,t}). \tag{3}$$

to refer to the MRTS variation driven by the adopted technology (at given input levels) and

$$\Delta\mathbb{F}^{s,u}_{cj,tT} = \left(MRTS^{s,u}_{cj,T} - MRTS^{s,u}_{cj,t}\right) \Big| \, (m_{cj,T} = m_{cj,t}). \tag{4}$$

to denote the MRTS variation driven by factor accumulation (and not by technology). The first component ($\Delta\mathbb{B}^{s,u}_{cj,tT}$) identifies the SBTC occurred from t to T in country c - sector j. Here we use the notation t to signify the initial period (in our application this will be 1995) and T to signify the final period (in our application this will be 2005). An alternative would be to use 0 and 1 or t_0 and t_1.

When both the skilled and unskilled labor markets are perfectly competitive, standard profit maximization yields the efficiency condition according to which the MRTS between skilled and unskilled labor (i.e., $MRTS^{s,u}_{cj,t} = \frac{\partial Y_{cj,t}/\partial S_{cj,t}}{\partial Y_{cj,t}/\partial U_{cj,t}}$) equals the wage ratio (i.e., $\mathbb{W}^{s,u}_{cj,t} = \frac{W^s_{cj,t}}{W^u_{cj,t}}$). Any deviation from this condition can be attributed to labor market imperfections affecting skilled and unskilled workers asymmetrically (see Caselli, 1999). Thus, a relative measure of economic inefficiency is given by the ratio

$$\tau^{s,u}_{cj,t} = \frac{MRTS^{s,u}_{cj,t}}{\mathbb{W}^{s,u}_{cj,t}}. \tag{5}$$

with $\tau^{s,u}_{cj,t} = 1$ denoting the absence of imperfections.

To see the relationship between SBTC and economic inefficiency, note that SBTC affects MRTS (i.e., the $\Delta\mathbb{B}^{s,u}_{cj,tT}$ term in (3)). To the extent that the MRTS variation induced by SBTC is not compensated by changes in relative wages and/or quantity of inputs (i.e., the $\Delta\mathbb{F}^{s,u}_{cj,tT}$ term in (4)), SBTC can be associated to a less efficient use of labor.

2.1 Some Motivation

Although we abstract from any type of functional/parametric specification in subsequent analysis, let us illustrate the intuition behind the identification of SBTC in a Cobb-Douglas setting, with a nested CES specification for S and U

$$Y_{cj,t} = Z_{cj,t}(K_{cj,t})^{\alpha}\left[(A^s_{cj,t}S_{cj,t})^{\sigma} + (A^u_{cj,t}U_{cj,t})^{\sigma}\right]^{\frac{1-\alpha}{\sigma}}, \tag{6}$$

Here Y denotes output in level and K, S and U refer to capital, skilled labor and unskilled labor in levels, respectively. $0 < \alpha < 1$ and $0 < \sigma \leq 1$, with the elasticity of substitution (EoS) between S and U given by EoS $= 1/(1-\sigma)$.

SBTC encompasses any form of technical change directly affecting the MRTS between skilled and unskilled labor, at given input levels. The MRTS, defined as the ratio of corresponding MPs (i.e., $\frac{\partial Y_{cj,t}/\partial S_{cj,t}}{\partial Y_{cj,t}/\partial U_{cj,t}}$), can be expressed as

$$MRTS^{s,u}_{cj,t} \equiv \frac{\partial Y_{cj,t}/\partial S_{cj,t}}{\partial Y_{cj,t}/\partial U_{cj,t}} = \left(\frac{A^s_{cj,t}}{A^u_{cj,t}}\right)^{\sigma}\left(\frac{S_{cj,t}}{U_{cj,t}}\right)^{\sigma-1} \tag{7}$$

with

$$\frac{\partial Y_{cj,t}}{\partial S_{cj,t}} = \Phi_{cj,t}(A^s_{cj,t})^{\sigma}(S_{cj,t})^{\sigma-1} \quad \text{and} \quad \frac{\partial Y_{cj,t}}{\partial U_{cj,t}} = \Phi_{cj,t}(A^u_{cj,t})^{\sigma}(U_{cj,t})^{\sigma-1} \tag{8}$$

and $\Phi_{cj,t} = (1-\alpha)Z_{cj,t}(K_{cj,t})^{\alpha}\left[(A^s_{cj,t}S_{cj,t})^{\sigma} + (A^u_{cj,t}U_{cj,t})^{\sigma}\right]^{\frac{1-\alpha-\sigma}{\sigma}}$.

Equation (7) highlights how MRTS varies according to changes in the labor ratio $\left(\frac{S_{cj,t}}{U_{cj,t}}\right)$, the EoS, and the ratio $\frac{A^s_{cj,t}}{A^u_{cj,t}}$, usually referred to as "relative efficiency". Given the first two, MRTS can only change because of relative efficiency: this is how SBTC is usually understood. Hence, the SBTC from t to T can be defined as the change in relative efficiency occurred during the period and expressed as

$$\Delta B_{cj,tT} = B_{cj,T} - B_{cj,t} \tag{9}$$

with

$$B_{cj,T} \equiv \frac{A^s_{cj,T}}{A^u_{cj,T}} = \left(MRTS^{s,u}_{cj,T}\right)^{\frac{1}{\sigma}} \left(\frac{S_{cj,T}}{U_{cj,T}}\right)^{\frac{1-\sigma}{\sigma}} \quad \text{and} \quad B_{cj,t} \equiv \frac{A^s_{cj,t}}{A^u_{cj,t}} = \left(MRTS^{s,u}_{cj,t}\right)^{\frac{1}{\sigma}} \left(\frac{S_{cj,t}}{U_{cj,t}}\right)^{\frac{1-\sigma}{\sigma}}. \tag{10}$$

In this parametric framework, SBTC can be measured only provided that reliable MRTS, labor ratio and EoS figures are available. As regards to MRTS, one would like to obtain it through estimating the MPs directly from (8). However, in an aggregate country-sector-year setting, this would entail estimating a model oversaturated with parameters unless various homogeneity restrictions were imposed (such as common technology across sectors or across countries). For such reason, the literature on SBTC has mainly followed the idea that the MRTS in Equation (10) can be inferred from relative wages, under the assumption that labor markets are perfectly competitive (Acemoglu and Autor, 2011). It is useful to illustrate the basics of this "indirect" (wage-based) approach before going into the details of our "direct" (production-based) approach working with Equation (3).

3 "Indirect" (wage-based) approach: intuition and shortcomings

The wage-based approach builds on the parametric specification in (10) under the hypothesis of perfectly competitive labor markets. Hence, we have

$$\frac{W^s_{cj,t}}{W^u_{cj,t}} = \left(\frac{A^s_{cj,t}}{A^u_{cj,t}}\right)^{\sigma} \left(\frac{S_{cj,t}}{U_{cj,t}}\right)^{\sigma-1} \equiv MRTS^{s,u}_{cj,t}, \tag{11}$$

where W^s and W^u are used to refer to the wage of skilled and unskilled labor, respectively. If the relative supply of skilled and unskilled labor is exogenous, relative wages fully describe the actual evolution of SBTC. Thus, the SBTC in (9) can be obtained residually from the change in the efficiency ratio from t to T, with the efficiency ratio obtained by using the wage ratio to replace for the MRTS in Equation (10):

$$B_{cj,t} \equiv \frac{A^s_{cj,t}}{A^u_{cj,t}} = \left(\frac{W^s_{cj,t}}{W^u_{cj,t}}\right)^{\frac{1}{\sigma}} \left(\frac{S_{cj,t}}{U_{cj,t}}\right)^{\frac{1-\sigma}{\sigma}}. \tag{12}$$

This intuition, carefully described in Caselli (2016), has been used in several studies at the firm or industry level (Katz and Murphy, 1992; Krusell et al., 2000; Card and Lemieux, 2001; Caselli and Coleman, 2006; Henderson, 2009; Violante, 2016 and Rossi, 2019).[6] While estimating SBTC in this fashion requires data on relative supply of skills, skill premium and EoS,[7] such an indirect approach

[6]Some authors (Koeniger and Leonardi, 2007; Alesina et al., 2018, among others) highlight how adopting a new technology can be relatively more or less convenient, depending on the pressure exerted by labor market institutions on wages. If this is the case, technology adoption is not independent of relative wages (Karabarbounis and Neiman, 2014). In other words, SBTC can be affected by the wage ratio. This emphasizes the importance of estimating SBTC directly from the production function, as suggested by our approach.

[7]This makes the approach both attractive and difficult to implement (see Caselli, 2016). In particular, it is worth stressing how SBTC measures obtained through the indirect approach are highly dependent on the choice concerning EoS

suffers from identification issues that we hereby discuss in brief.

"Biased" biased technical change. By relying on the assumption that the whole deviation from the efficiency condition is due to technical change, the approach in (12) and (9) attributes to SBTC the effect of any type of eventual frictions affecting the market of skilled and unskilled labor differently (see also the discussion in Card and Di Nardo, 2002 and Caselli, 2016). To see this using Equation (11), consider how an increase in $\frac{A^s_{cj,t}}{A^u_{cj,t}}$ results into an increasing relative demand of skilled labor (in order to compensate the increase in MRST), which in turn pushes towards an increasing wage gap in the aggregate.

In a perfectly competitive labor market, these two effects generate a new economically efficient situation in which SBTC is associated with both a higher $\frac{W^s_{cj,t}}{W^u_{cj,t}}$ and a higher $\frac{S_{cj,t}}{U_{cj,t}}$. The presence of asymmetric imperfections affecting skilled labor differently from unskilled labor can prevent such an adjustment from taking place, thereby opening a "wedge" between MRTS and relative wages. This economically inefficient situation can be visualized using parameter $\tau^{s,u}_{cj,t} > 1$ to subsume such wedge. With frictions resulting into an inefficiently low realized wage ratio $\frac{W^s_{cj,t}}{W^u_{cj,t}}$, compared to the perfectly competitive one – i.e., $\tau^{s,u}_{cj,t}\frac{W^s_{cj,t}}{W^u_{cj,t}}$, we have:

$$MRTS^{s,u}_{cj,t} = \left(\frac{A^s_{cj,t}}{A^u_{cj,t}}\right)^{\sigma} \left(\frac{S_{cj,t}}{U_{cj,t}}\right)^{\sigma-1} = \tau^{s,u}_{cj,t}\frac{W^s_{cj,t}}{W^u_{cj,t}}. \tag{13}$$

Equation (13) yields the following expression for the "true" relative efficiency term

$$B^*_{cj,t} = \tau^{s,u}_{cj,t}\underbrace{\left(\frac{W^s_{cj,t}}{W^u_{cj,t}}\right)^{\frac{1}{\sigma}} \left(\frac{S_{cj,t}}{U_{cj,t}}\right)^{\frac{1-\sigma}{\sigma}}}_{B_{cj,t}} \tag{14}$$

where $B_{cj,t}$ is the (biased) relative efficiency term in (12), computed without taking labor market inefficiency into account.

Hence, the true SBTC can be written as

$$\Delta B^*_{cj,tT} = (\tau^{s,u}_{cj,T})^{\frac{1}{\sigma}} B_{cj,T} - (\tau^{s,u}_{cj,t})^{\frac{1}{\sigma}} B_{cj,t} \tag{15}$$

Equation (15), which nests (9) as a special case with $\tau^{s,u}_{cj,T} = \tau^{s,u}_{cj,t} = 1$, highlights how the indirect approach conflates "true" SBTC and labor market vicissitudes in the presence of asymmetric distortions between skilled and unskilled labor markets. In particular, SBTC is understated whenever $\tau^{s,u}_{cj,t} > 1$, that is in presence of distortions compressing the wage of skilled labor relative to unskilled labor (the opposite case is discussed in Caselli and Coleman, 2006). While the understatement grows with the asymmetry[8] (that is, if $\tau^{s,u}_{cj,T} > \tau^{s,u}_{cj,t}$), the usual assumption in the indirect approach is the absence of asymmetric labor

parameter. In principle, the EoS can be obtained, as suggested by Equation (11), by regressing relative wages on labor ratios, usually at the country level including controls (Katz and Murphy, 1992; Acemoglu and Autor, 2011).

[8] A time-invariant asymmetry $\tau^{s,u}_c = \tau^{s,u}_{cj,T} = \tau^{s,u}_{cj,t}$ does not eliminate the understatement, as we have

$$\Delta B^*_{cj,tT} = (\tau^{s,u}_c)^{\frac{1}{\sigma}}\left[B_{cj,T} - B_{cj,t}\right] = (\tau^{s,u}_c)^{\frac{1}{\sigma}}\,\Delta B_{cj,tT}.$$

market frictions (i.e., $\tau^{s,u}_{cj,T} = \tau^{s,u}_{cj,t} = 1$), which eventually boils down to assuming perfect competition in both the skilled and unskilled labor markets.

An *ad hoc* specification. In the indirect approach, SBTC is identified only under very specific assumptions about how S and U enter the production function (i.e., the CES hypothesis). To see this, consider the alternative specification

$$Y_{cj,t} = Z_{cj,t}(K_{cj,t})^{\alpha} \left[(A^s_{cj,t}S_{cj,t})^{\sigma}(A^u_{cj,t}U_{cj,t})^{1-\sigma}\right]^{1-\alpha}. \tag{16}$$

The derivatives with respect to S and U are given by, respectively:

$$\frac{\partial Y_{cj,t}}{\partial S_{cj,t}} = \sigma(1-\alpha)\frac{Y_{cj,t}}{S_{cj,t}} \qquad \text{and} \qquad \frac{\partial Y_{cj,t}}{\partial U_{cj,t}} = (1-\sigma)(1-\alpha)\frac{Y_{cj,t}}{U_{cj,t}} \tag{17}$$

with the MRTS between skilled and unskilled labor amounting to

$$\frac{\partial Y_{cj,t}/\partial S_{cj,t}}{\partial Y_{cj,t}/\partial U_{cj,t}} = \frac{\sigma}{1-\sigma}\frac{U_{cj,t}}{S_{cj,t}}. \tag{18}$$

In this case, the relative efficiency term $\frac{A^s_{cj,T}}{A^u_{cj,T}}$ disappears from the MRTS, and as does SBTC, as long as the EoS parameter is fixed across countries, sectors and time (see discussion in next paragraph).

σ-dependence. From Equation (12), SBTC is identified up to the scale based on σ, which encapsulates the EoS between skilled and unskilled labor. If σ does not vary across countries, sectors or time, such a scale effect can be easily removed by either using an estimated value from the empirical literature (e.g., Autor et al., 2008 suggest values ranging from 1 to 2; Caselli, 2016 assumes a value of 1.5, entailing $\sigma = 1/3$) or expressing the SBTC in relative terms with respect to a benchmark (e.g. the US as in Caselli, 2016). However, the empirical literature has largely documented how the EoS can differ across industries (Krusell et al., 2000; Blankenau and Cassou, 2011), countries and time (Henderson, 2009). Moreover, the specification described in the previous paragraph shows how a country-sector-time invariant EoS can easily make SBTC unidentifiable (see Equation (18)).

More realistically, σ can be thought of as a technological parameter itself varying across countries, sectors and time in response to the paths of technological change characterizing different countries and sectors (i.e., $\sigma_{cj,t}$). This assumption yields different effects in the functional specifications (6) and (16). In the former, $\sigma_{cj,t}$ remains a scale parameter conflating the magnitude of SBTC but its (country-sector-time) heterogeneity makes its removal from these figures problematic (neither using exogenous values, as long as country-sector-time specific figures are not available, nor switching to relative measures). In the latter, heterogeneity in σ affects the MP of S and U in different proportions and, as such, enters SBTC directly – i.e., $B_{cj,t} \equiv \frac{\sigma_{cj,t}}{1-\sigma_{cj,t}} = \left(\frac{W^s_{cj,t}}{W^u_{cj,t}}\right)\left(\frac{S_{cj,t}}{U_{cj,t}}\right)$. In this case, SBTC is easily identified but the EoS is the driver of SBTC, as the efficiency ratio cancels out.

"Mechanical" prediction. Equation (12) "mechanically" determines the direction of the bias associated with technical change. In particular, in the presence of growing wage premia, SBTC can only arise if changes in relative wages and changes in labor ratios are not inversely related (i.e., if labor ratios also grow). By contrast, with decreasing wage ratios, technical change is skill biased only if the labor ratio grows in such a way that the growth in $\left(\frac{S_{cj,t}}{U_{cj,t}}\right)^{1-\sigma}$ more than compensates the decrease in $\left(\frac{W^s_{cj,t}}{W^u_{cj,t}}\right)$, entailing an inverse relationship between changes in relative wages and changes in labor ratios. Otherwise, technological progress is biased in favor of unskilled labor. Moreover, the higher the EoS (σ), the higher is the required increase in the labor ratios. As Card and Di Nardo (2002) emphasize, this is a fuzzy implication in all those cases in which the wage premium fails to grow at the pace of technology adoption (as in the US in the 1990s).

As shown in Figure 1, the wage ratio decreased in many countries over the 1995-2005 decade, sometimes considerably (Italy, Sweden, Brazil). In those cases, the indirect approach would likely predict technological progress to be biased in favor of unskilled labor, to the extent that the increase in the labor ratio is relatively low and the EoS is relatively high. This raises questions about whether indirect SBTC figures are indeed capturing the true essence of SBTC notwithstanding the presence of (potentially asymmetric between skilled and unskilled) labor market imperfections.[9]

4 A "direct" (production-based) approach

As said, instead of relying on the wage ratio to proxy for MRTS in Equation (10), one might want to estimate MRTS directly. While this is the most intuitive way to deal with SBTC, this approach has not been followed so far because of the econometric issues associated with estimating the MP of each input in a given country-sector-year. In a standard parametric setup this approach is problematic as the number of parameters to estimate would easily exceed the number of observations. To overcome this limitation, we suggest a nonparametric approach.

Identification relies on i) estimating the MRTS between skilled and unskilled labor; ii) using the estimated MRTS to obtain the SBTC in (3) through counterfactual analysis.

4.1 Estimating MRTS

The first step in the analysis is the estimation of the MRTS between skilled and unskilled labor from (1) allowing for cross-country, cross-sector and cross-time variation. For this, we treat the production function in (1) as an unknown smooth function varying across the three dimensions (Battisti et al., 2018) and use local-linear least-squares (LLLS), which performs weighted least-squares around a given point, with weights determined by a kernel function and a bandwidth vector (more weight is given

[9] Baqaee and Farhi (2020) recently developed a general theory of aggregation from the firm level in inefficient economies. Using a nonparametric decomposition of the changes in aggregate TFP into pure (exogenous) changes in technology and (endogenous) changes in allocative efficiency, they find that improvement in allocative efficiency, through reallocation of market share to high-markup firms over time, would account for about half the aggregate TFP growth over the period 1997-2015.

to observations in the neighborhood). LLLS allow smoothing over continuous (k, s and u), ordered (time) and unordered (country-sector) covariates simultaneously. The ability to smooth over discrete cells affords us the ability to estimate time and country-sector specific MPs.

Specifically, we write $m_{cj,t}(\cdot)$ in Equation (1) as

$$y_{cj,t} = m(\boldsymbol{x}_{cj,t}, d_{cj}, d_t), \tag{19}$$

where, aside from the continuous variables $k_{cj,t}$, $s_{cj,t}$ and $u_{cj,t}$, we use two discrete variables: a country-sector effect d_{cj} (which is constant over time) and a time effect d_t. While the latter is ordered by nature, the former has no natural ordering. In a parametric setting, allowing for country-sector-year effects would introduce a large and infeasible number of unobserved effects; moreover, even accounting for country-sector-year effects would not allow for heterogeneity of marginal products of production unless there were also interactions that were included in the model, quickly eliminating degrees of freedom. By smoothing across both time and sector, we can lessen the impact of common parametric strategies (such as time and country-sector intercept shifts) by leveraging "nearby" cells for local information.[10] The ability to smooth discrete cells allows local averaging of nearby cells to provide more "observations" to measure the production function at a point. As the sample size increases, this necessarily allows the neighborhood size to be reduced.

To make the notation a bit more manageable, let us drop the sector by country denomination and define our continuous variables as $\boldsymbol{x}_t = (\boldsymbol{x}_{cj,t}) = k_{cj,t}, s_{cj,t}, u_{cj,t}$ and our discrete variables as $\boldsymbol{d} = (d_{cj}, d_t)$. We estimate the model by kernel smoothing a local-linear approximation of (1) via a first-order Taylor expansion around a given point $\boldsymbol{x}$ (note that the expansion is only for the three continuous covariates k, s, and u) according to

$$\begin{aligned} y_t &= m(\boldsymbol{x}_t, \boldsymbol{d}) + \varepsilon_t \\ &\approx m(\boldsymbol{x}, \boldsymbol{d}) + (\boldsymbol{x}_t - \boldsymbol{x})\beta(\boldsymbol{x}, \boldsymbol{d}) + \varepsilon_t \end{aligned} \tag{20}$$

where $m(\boldsymbol{x}_t, \boldsymbol{d}) = m_{cj,t}(\boldsymbol{x}_{cj,t}, d_{cj}, d_t)$ and $(\boldsymbol{x}_t - \boldsymbol{x})'$ is a 3×1 vector. The vector $\beta(\boldsymbol{x}, \boldsymbol{d})$ is defined as the partial derivative vector of $m(\boldsymbol{x}_t, \boldsymbol{d})$ with respect to $\boldsymbol{x}$ and is an estimate of the three MPs: $\frac{\partial m_{cj,t}}{\partial k_{cj,t}}$, $\frac{\partial m_{cj,t}}{\partial s_{cj,t}}$, and $\frac{\partial m_{cj,t}}{\partial u_{cj,t}}$ (which we denote as $m^k(\cdot)$, $m^s(\cdot)$ and $m^u(\cdot)$, respectively).[11]

Appendix A contains more details on both the estimation and bandwidth selection method.

The generalized kernel regression approach is particularly appealing to our purposes because of its

[10] This comes at the expense of introducing bias into the estimators (Li and Racine, 2007) but has the potential to lower variance and has been shown to lead to substantial finite sample gains (Li and Racine, 2004). Li and Racine (2004) focus on the iid setting, whereas our setting here is more aptly characterized by data dependence over time. Robinson (1983) (for strongly dependent data), Li and Racine (2007, chapter 18), (for a martingale difference process), and Li et al. (2009) (for weakly dependent mixed discrete and continuous data), demonstrate that the main large sample properties of the regression estimator in the iid setting carry over to the dependent data case. In our back of the envelope investigation, we deploy several alternative data checks to determine if various forms of time series dependence may have undue influence on our results. These robustness checks are available upon request.

[11] In the CES parametric setup described above, Equation (20) would read, using index i to refer to the generic cj, t

ability to estimate $m(\boldsymbol{x}, \boldsymbol{d})$ and $\beta(\boldsymbol{x}, \boldsymbol{d})$ including country-sector and year effects. As discussed earlier, this would be difficult in a standard parametric setup, as the number of parameters to be estimated could easily surpass the number of observations (country c at time t in sector j). By smoothing over continuous, ordered and unordered covariates simultaneously, our approach allows us to identify a country-sector-year specific relationship between output and inputs, thereby differing from conventional estimates that restricts the estimated relationship to have a common form across countries (input coefficients are sector specific but not country specific).

As Li and Racine (2004) and Li et al. (2009) show, the addition of discrete regressors does not affect the convergence rate of the conditional mean, as it is *only* dependent on the number of continuous regressors. This is quite important, as the curse of dimensionality is one of the primary criticisms against the use of nonparametric methods in empirical studies. In our application, with two of the five variables being discrete, our sample size should be adequate to learn about the underlying production technologies.

Hence, our nonparametric approach allows us to estimate country-sector-time specific partial derivatives for each input (see Appendix A for details) avoiding explicit assumptions on the functional form of technology (e.g. Cobb-Douglas). We only require that log-TFP enters additively, ending up in $z_{cj,t}$.

Following Equation (2), the MRTS can be obtained as the ratio of the estimated partial derivative with respect to S to the estimated partial derivative with respect to U:

$$\widehat{MRTS}^{s,u}_{cj,t} = \frac{\widehat{m}^s_{cj,t}(\boldsymbol{x}_{cj,t})}{\widehat{m}^u_{cj,t}(\boldsymbol{x}_{cj,t})}. \tag{21}$$

It is worth noting how the estimated MRTS in (21) is virtually net of any form of Hicks-neutral technical change, as the log-additive term $z_{cj,t}$ vanishes when the partial derivative is estimated, while eventual log-multiplicative components are swept away when computing the ratio of the estimated MP.

The rate of change of MRTS can be defined as:

$$\Delta \ln \widehat{MRTS}^{s,u}_{cj,tT} = \ln\left(\frac{\widehat{MRTS}^{s,u}_{cj,T}}{\widehat{MRTS}^{s,u}_{cj,t}}\right) = \ln\left(\frac{\widehat{m}^s_{cj,T}(\boldsymbol{x}_{cj,T})}{\widehat{m}^u_{cj,T}(\boldsymbol{x}_{cj,T})}\right) - \ln\left(\frac{\widehat{m}^s_{cj,t}(\boldsymbol{x}_{cj,t})}{\widehat{m}^u_{cj,t}(\boldsymbol{x}_{cj,t})}\right). \tag{22}$$

Note that MRTS changes can depend on either changes in technology, through SBTC, or changes in the amount of inputs used (see Equation 8 for a standard parametric specification). The latter change should be isolated otherwise it may be a driving force in affecting results (see for instance Buera et al, 2021). Thus, we hereby proceed with disentangling the two effects through counterfactual analysis.

observation, as:

$$y_i \approx m_i(x_i) + (k_i - k)\underbrace{\left[\ln\alpha + \ln y_i - \ln k_i\right]}_{\beta_i^k} + (s_i - s)\underbrace{\left[\ln(\Phi_i) + \ln((A_i^s)^\sigma S_i^{\sigma-1})\right]}_{\beta_i^s} + \\ + (u_i - u)\underbrace{\left[\ln(\Phi_i) + \ln((A_i^u)^\sigma U_i^{\sigma-1})\right]}_{\beta_i^u} + \varepsilon_i$$

where (8) has been used and $m_i(x_i)$ is defined by Equation (6).

4.2 Dissecting MRTS growth: SBTC and FA

To isolate the SBTC component from the MRTS change, as suggested by Equation (3), a counterfactual analysis is necessary, in which we measure the change in output (associated with technological progress) that the country would have experienced without changing the amount of inputs used. Indeed, for each input (say skilled labor) we are able to compute counterfactual partial derivatives $\widetilde{m}^{s}_{cj,T}(\boldsymbol{x}_{cj,t})$ measuring the additional output (associated to the marginal unit of S) that country c would have produced at time t using time T's technology, given the quantity of inputs used.

These counterfactual values can be used to decompose the MRTS variation in (22) into the following two components:

$$\Delta \ln \widehat{MRTS}^{s,u}_{cj,tT} = \Delta \ln \widetilde{\mathbb{B}}^{s,u}_{cj,tT} + \Delta \ln \widetilde{\mathbb{F}}^{s,u}_{cj,tT}. \tag{23}$$

The first term, defined as

$$\Delta \ln \widetilde{\mathbb{B}}^{s,u}_{cj,tT} = \ln \widetilde{MRTS}^{s,u}_{cj,T} - \ln \widehat{MRTS}^{s,u}_{cj,t} = \ln \left(\frac{\widetilde{m}^{s}_{cj,T}(\boldsymbol{x}_{cj,t})}{\widetilde{m}^{u}_{cj,T}(\boldsymbol{x}_{cj,t})} \right) - \ln \left(\frac{\widehat{m}^{s}_{cj,t}(\boldsymbol{x}_{cj,t})}{\widehat{m}^{u}_{cj,t}(\boldsymbol{x}_{cj,t})} \right), \tag{24}$$

identifies SBTC insofar it encompasses the technical change not affecting the marginal revenue product of skilled and unskilled labor in the same proportion.

The second term, defined as

$$\Delta \ln \widetilde{\mathbb{F}}^{s,u}_{cj,tT} = \ln \widehat{MRTS}^{s,u}_{cj,T} - \ln \widetilde{MRTS}^{s,u}_{cj,T} = \ln \left(\frac{\widehat{m}^{s}_{cj,T}(\boldsymbol{x}_{cj,T})}{\widehat{m}^{u}_{cj,T}(\boldsymbol{x}_{cj,T})} \right) - \ln \left(\frac{\widetilde{m}^{s}_{cj,T}(\boldsymbol{x}_{cj,t})}{\widetilde{m}^{u}_{cj,T}(\boldsymbol{x}_{cj,t})} \right), \tag{25}$$

is a FA component measuring the MRTS variation related to the change in the quantity of inputs (factor accumulation), at a given level of technology.

4.3 Measuring economic inefficiency

This approach nicely fits into an economic efficiency analysis. As highlighted in Section 2, under standard profit maximization, the MRTS equals the wage ratio:

$$MRTS^{s,u}_{cj,t} = \mathbb{W}^{s,u}_{cj,t} = \frac{W^{s}_{cj,t}}{W^{u}_{cj,t}}. \tag{26}$$

Thus, with the estimated MRTS in our hands, the presence of market imperfections affecting skilled and unskilled labor asymmetrically can be evaluated by computing the relative measure of economic inefficiency in equation (5), which also approximates the term $\tau^{s,u}_{cj,t}$ in Equation (13):

$$\ln \widehat{\tau}^{s,u}_{cj,t} = \ln \left(\frac{\widehat{MRTS}^{s,u}_{cj,t}}{\mathbb{W}^{s,u}_{cj,t}} \right). \tag{27}$$

The efficiency condition implies $\tau^{s,u}_{cj,t} = 1$. When $\tau^{s,u}_{cj,t} > 1$, the MP of skilled labor, relative to unskilled labor, is too high, compared to the actual wage ratio. The opposite is true for $\tau^{s,u}_{cj,t} < 1$.

Since we estimate the MRTS in Equation (21) at different points in time, we are able to relate the variation in $\widehat{\tau}^{s,u}_{c}$ to SBTC as follows:

$$\Delta \ln \widehat{\tau}^{s,u}_{cj,tT} = \ln \widehat{\tau}^{s,u}_{cj,T} - \ln \widehat{\tau}^{s,u}_{cj,t} = \underbrace{\Delta \ln \widetilde{\mathbb{B}}^{s,u}_{cj,tT} + \Delta \ln \widetilde{\mathbb{F}}^{s,u}_{cj,tT}}_{\ln\left(\frac{\widehat{MRTS}^{s,u}_{cj,T}}{\widehat{MRTS}^{s,u}_{cj,t}}\right) = \Delta \ln \widehat{MRTS}^{s,u}_{cj,tT}} - \Delta \ln \mathbb{W}^{s,u}_{cj,tT} \tag{28}$$

where $\Delta \ln \mathbb{W}^{s,u}_{cj,tT} = \ln\left(\frac{\mathbb{W}^{s,u}_{cj,T}}{\mathbb{W}^{s,u}_{cj,t}}\right)$. Equation (28) provides information concerning the eventual decoupling between relative productivities and relative wages: positive values imply that the MP of skilled labor has grown relatively more than wages, compared to unskilled labor. Moreover, we are able to determine the extent to which such decoupling is driven by SBTC, rather than by the amount of inputs used.

Notably, the extent of economic inefficiency grows with the distance of $\tau^{s,u}_{c}$ from zero, no matter whether values are positive or negative. Thus, it is convenient to express the variation of $\widehat{\tau}^{s,u}_{c}$ using absolute values to operationalize the formulation in Equation (28) in terms of "increasing" or "decreasing" economic inefficiency as follows:

$$\widehat{\Theta}^{s,u}_{cj,tT} = \left|\ln \widehat{\tau}^{s,u}_{cj,T}\right| - \left|\ln \widehat{\tau}^{s,u}_{cj,t}\right|. \tag{29}$$

$\widehat{\Theta}^{s,u}_{cj,tT}$ measures the change in economic efficiency in relative terms. Positive values point to increasing economic inefficiency, a widening discrepancy between relative wages and MRTS. This circumstance can be driven by a change in relative wages, the adoption of skill biased technologies (i.e., SBTC), a change in the skilled to unskilled labor actually used (i.e., factor accumulation).[12] The absence of changes in economic efficiency (i.e., $\widehat{\Theta}^{s,u}_{cj,tT} = 0$) does not rule out SBTC.

To sum up, the basic idea of the direct approach described in this Section consists of estimating the MPs of inputs (i.e., the partial derivatives) directly from the production function and using them to obtain the decomposition in Equation (23). This entails three advantages over the indirect approach described in Section 3. First, estimation only requires information on hired quantities of labor (i.e., equilibrium employment by skill types); wages are not needed. Second, given that SBTC is obtained independently of the observed wage gap, our estimates are not affected by the presence of labor market imperfections. Indeed, the evolution of the discrepancy between relative wages and MRTS (referred to as economic inefficiency) can be studied ex post. Third, our approach does not require an estimate of σ and is not affected by the scale issues highlighted in Section 3. In particular, the nonparametric estimation ensures that the technological progress affecting S and U in different proportions is fully captured by the SBTC component in (23) (i.e., $\Delta \ln \widetilde{\mathbb{B}}^{s,u}_{cj,tT}$), with no σ-dependance troubling the estimates and with the EoS flowing into the estimated SBTC as long as it represents a technological parameter whose country-

[12] Although rare, positive values of $\widehat{\tau}^{s,u}_{cj,T}$ sometimes turn into negative values, and vice-versa. Since the economic inefficiency is violated in both cases, absolute values are needed. This issue is, to some extent, similar to the β-convergence problem in the presence of leap frogging effects.

sector-time heterogeneity can be interpreted in terms of differential effect of technological progress on the MP of S and U.

5 SBTC analysis

5.1 Data

In this section we estimate the MRTS and bring the decomposition in Equation (23) to data, basing the analysis on the Socio-Economics Accounts section of WIOD (World Input Output Database; Timmer *et al.*, 2015; Erumbam *et al.*, 2012). This database contains detailed data on annual basis for 40 countries and 35 (ISIC Rev. 3) sectors, from 1995 to 2009 (release February 2012). Values are expresssed at 1995 constant prices. Specifically, we use:

- Value added (in current value) as the output measure Y;
- Real fixed capital stock as K;
- High, medium and low-skilled labor, expressed in terms of employment by category, obtained by multiplying the total employment in the country by the share of hours worked in the single category.[13];
- Hourly wages for high, medium and low-skilled labor. Hourly wages are computed as the ratio of share in labor compensation times total labor compensation to share in hours worked times total hours worked.

To reconcile this three categories with our two-categories specification encompassing only S and U, we include medium-skilled labor in U, in the benchmark specification, and in S, in the robustness section. Current price variables are converted into real terms using the price level of gross value added and then transformed them into PPP using absolute 1995 PPPs variable XR. [14] We undertake benchmark analysis over the averages of two periods, 1995-1996 and 2005-2006; this allows us to remain somehow removed with the effect of the economic crisis. After cleaning for missing, zero and negative value added values, we are left with a final sample of 975 observations at both the beginning and the end of the period, covering 38 countries and 25 sectors.

Tables 1 and 2 report the (by country and by sector, respectively) descriptive statistics at the beginning and the end of the period, highlighting an average 3.7% increase in value added, associated with a 4.5% increase in capital. The stock of skilled labor displays a huge increase (11.5%), while unskilled labor decreases by -0.6%.

[13]WIOD provides skill shares at the industry level. For EU countries, for instance, these come from the Labor Force Survey.

[14]Absolute 1995 PPPs are computed as exchange rate from Penn World Tables 9.1 (Feenstra *et al.*, 2015) times relative PPP at 1995.

5.2 Benchmark estimation

As a first step we estimate the $\beta(x, d)$ vector (i.e., the MPs) in Equation (20) as described in Section 4.1. The estimated distribution is discussed in Appendix C.

The MRTS between skilled and unskilled labor at the beginning and the end of the period under consideration is then obtained, for each country-sector, as the ratio in (21). These values are summarized in Table 3, together with standard errors, with the overall distribution and confidence intervals plotted in Figure 2. The third row of the table reports the counterfactual value of MRTS computed with technology of 2005 and input stocks of 1995.

The estimated MRTS changes are detailed in Table 4, by country, and Table 5, by sector (unweighed averages reported).[15] Overall, we report a positive change in MRTS (around 11.5% on average), which, given our ten-year time span, maps into a roughly 1.1% yearly increase in the MP of skilled labor, relative to unskilled workers. Within this average figure, we observe substantial heterogeneity.

To understand the extent the change in terms of economic efficiency can be traced back to SBTC, we bring the counterfactual decomposition presented in Section 4 to data by computing Equation (23). Results are detailed by country in Table 4, and by sector in Table 5.

The driving force of the MRTS variation is SBTC, which is 13.7% on average, that is around 1.37% on a yearly basis. Although with a certain degree of heterogeneity across sectors and countries, the average role of FA is negative and relatively small (around 2% on average). In addition, we observe substantial heterogeneity across countries and sectors. The FA effect sometimes dominates the SBTC component. Notably, this is the case in the US, for which we estimate a -19.5% decease in the MRTS associated with a huge negative FA contribution (-48.6%) and a positive SBTC (29%). We detect positive SBTC terms for all sectors, ranging from the lowest in retail trade to the highest in basic and fabricated metals (out of real estate that has too few observations). The FA contribution is negative in several industries and countries. According to our framework, this suggests that keeping the technology in use constant, the evolution of the (S,U) input mix is such that the overall contribution to the MRTS growth rate is negative; this might be interpreted in terms of decreasing MPs, as long as the relative use of S has been growing over the 1995-2005 decade (see Figure 1). By contrast, decreasing MRTS can always be traced back to a strongly negative FA effect.

An interesting dimension (see Figure 3) is that it is possible to identify a group of (mainly less developed) countries in which we observe a relatively high increase in the MRTS mainly driven by the FA dimension. In particular, this is the case of India, the third country in terms of MRTS growth, featuring an even negative SBTC. On the opposite side, we find countries like Korea, USA, Japan, Germany, in which the MRTS is even negative, mainly due to FA. In particular, while displaying the third highest SBTC figure (29%), USA features the second last decrease in MRTS (-19.5%). These examples might reveal an attitude to accumulate skilled labor, relative to unskilled labor, thereby decreasing the relative productivity of the former, when technical change is skill biased; and vice-versa.

[15] In unreported analysis we verified that figures are pretty similar when the country (sector) averages are weighted using the sectoral (country) shares within the country (sector).

5.3 Broad validation

The unavailability of comparable SBTC figures in the literature prevents us from carrying out a benchmarking analysis. This notwithstanding, it is important to get a sense of meaningfulness of our SBTC figures. To this aim, we present a broad validation exercise in which we ask to what extent our estimated SBTC correlates with country-sector measures of IT exposure.

We use the WIOD information on output to generate a Balassa-type index of sectoral specialization. We then use EUKLEMS[16] information on IT capital to compute world average and US-specific sectoral shares of ICT capital. These shares are interacted interacted with the index of sectoral specialization to obtain country-sector measures of IT exposure that we expect to be positively correlated with our estimated SBTC.

The regression output of SBTC against these two country-sector variables is reported in the first and second row of Table 6. In the third row we replace the IT capital variable with the sectoral robot stock at the world level (i.e., global value of installed robots in a given sector). In row four we replace the Balassa index with a Herfindal index of output concentration. All results point to a strong and positive correlation, as expected.

5.4 Robustness checks

In this section, we report several checks aimed at evaluating the robustness of our results. Table 7 reports these findings. For ease of comparison, we also report our benchmark results.

In our benchmark analysis, we use two-years beginning and end of period averages in order to smooth over yearly anomalies. As a first check (row 1), we replicate our estimates using 1995 and 2005 as the first and last year, without averaging. In row 2, we use three year averages centered in 1996 and 2006.

In row 3, we report the estimates obtained using a least-squares cross-validation (LSCV) procedure to select the bandwidths as opposed to our rule-of-thumb approach. Data-driven selection of bandwidths is viewed as more agnostic towards the level of smoothing but is also susceptible to optimal bandwidths which are empirically small, leading to noisy estimates which are potentially seen as unpalatable in applied work (Loader, 1999). Here our LSCV bandwidths produce similar conclusions to our benchmark rule-of-thumb bandwidths, which is reassuring that our results are not being driven entirely by smoothing parameter selection. Although the number of observations slightly changes across the different specifications (in particular, it shrinks by about 6-7% in the "Not averaged 1995-2005" case) because of time changing country-sector coverage, the results are very close to our benchmark figures.

As a further robustness check (row 4), we consider our measures when we restrict technology to be identical across countries within sectors.[17] Here this should set SBTC equal across countries within a sector (assuming similar levels of capital and labor). While this may not be a palatable assumption, it is a useful middle ground between our proposed nonparametric approach and the common parametric

[16] The EUKLEMS database (http://www.euklems.net/) provides country-sector information on capital (K), labour (L), energy (E), materials (M) and service (S) inputs for EU countries.

[17] We thank an anonymous referee for this suggestion.

approach that assumes this by default. Moreover, if one believes that countries should be exposed to similar production technologies this may be seen as a useful prediction. In this case our estimate of SBTC is larger than in our benchmark results, which aligns with intuition.

5.5 Potential limits of the direct approach

"High-Medium vs Low" skilled labor. While the WIOD database provides detailed information on three classes of labor skills (high, medium and low), we use a "High vs Medium-Low" classification in our benchmark analysis. Certainly it would be preferable to distinguish each of these three classes. We do not do this for two specific reasons. First, nonparametric kernel smoothing methods have increased finite sample bias as the number of continuous covariates is increased (the curse of dimensionality) and we already have a capital stock variable and the two broad classes of labor market skills. Second, our decomposition of MRTS and SBTC would require further modification with three labor classes. While these circumstances might be perceived as intrinsic limits of our nonparametric approach, our belief is that the essence of our focus is maintained with two labor market skill classes while also assisting in the interpretation of the analysis. However, the results might differ depending on the classification adopted. To address this issue, Table 7 (row 5) reports the results obtained classifying labor into the two categories of "High-Medium" and "Low". Overall, the sign of the differences accords with expectations. Indeed, the higher magnitude of the estimated MRTS and SBTC effects is expected and the insights with this different arrangement of labor market skills are broadly consistent with an increase, instead of a decrease, in relative wages associated with a higher SBTC and a substantially lower FA effect. Consequently, the estimated change in economic efficiency is higher.

Accounting for Endogeneity. In principle, estimation of (1) might suffer from potential endogeneity of $m_{cj,t}(\cdot)$ with respect to aggregate Hicks-neutral total factor productivity $z_{cj,t}$ (i.e., TFP), to the extent this is observed and plays a role in firms' choices about technology and amount of inputs.[18]

This "transmission bias" might distort our MRTS and SBTC estimates. However, it is worth noting that this is the case only provided that the transmission mechanism changes over time; otherwise, given that the analysis is entirely carried out in growth changes, it cancels out. Moreover, it is worth noting how, while a country-sector reaction to an aggregate productivity shock entails reallocating S and U across sectors and/or countries, cross-industry and (even more) cross-country mobility is usually low and tends to decline with sectoral aggregation (see, among others, Kambourov and Manovskii, 2008; Foster

[18]The endogeneity (or simultaneity) issue has been highlighted by the literature on firm-level production function estimation. The source of endogeneity consists is the fact that information on the firm's TFP, although unknown to the econometrician, is known to the firm when it decides on the amount of inputs. This biases the production function parameters obtained through OLS estimates because of the potential correlation between the regressors and the error term. A common assumption in the literature addressing the simultaneity bias is about productivity of firm i, operating in country c - sector s, to evolve according to a Markov process - i.e., $Z_{ics,t} = E[Z_{ics,t}|Z_{ics,t-1}] + \xi_{ics,t}$. In principle, the correlation of such an idiosyncratic shock with $m_{cj,t}(\cdot)$ and the vector of inputs should disappear at the country-sector level. However, also aggregate TFP might play a role in firm i's input and technology choice. This could be the case when the productivity "surprise" $\xi_{ics,t}$ consists of a idiosyncratic component ($\eta_{ics,t}$) and a country-sector component ($\mu_{cs,t}$), so that $Z_{ics,t} = E[Z_{ics,t}|Z_{ics,t-1}] + \underbrace{\mu_{cs,t} + \eta_{ics,t}}_{\xi_{ics,t}}$. In this case, $m_{cj,t}(\cdot)$ estimates would be biased by correlation with $\mu_{cs,t}$.

Mc Gregor and Poschl, 2009; Neffke et al., 2017; Park et al., 2019).

This notwithstanding, to address the potential endogeneity bias we hereby carry out an analysis in which we assume that time t's inputs are affected by observed TFP in period $t-1$.[19] Assuming that the TFP shock follows a first order Markov process $z_{cj,t} = h_{cj,t}\left(z_{cj,t-1}\right) + \varepsilon_{cj,t}$, we borrow from the insights of the nonparametric instrumental variable estimator of Su and Ullah (2008) to estimate a production function of the type $y_{cj,t} = m_{cj,t}(\boldsymbol{x}_{cj,t}) + h_{cj,t}(z_{cj,t-1}) + \varepsilon_{cj,t}$ by including recursively the estimated residuals obtained by estimating the same production function for the previous time period to proxy for $h_{cj,t}(z_{cj,t-1})$. While the estimation procedure is detailed in Appendix B, the output of such exercise, centered in years 1996 and 2006, is reported in row 6 of Table 7. SBTC grows substantially, up to 3.8%. The magnitude of the FA component points to a slightly larger negative contribution to the overall MRTS change (pointing for the presence of steeper relative decreasing marginal returns), which grows slightly less than in the benchmark analysis. Thus, it appears that even in the presence of endogeneity our main benchmark conclusions hold.

Skill quality change over time. It is worth noting how skill quality might be a confounding factor in MRTS estimation. However, according to extant literature, there seems to be little room for this. Indeed, Caselli (2005) finds no relevant role for quality of education in a development accounting framework and Rossi (2019) shows that technological change explains the variation of skilled labor efficiency much better than the change in human capital quality. For the US labor market, Carneiro and Lee (2011) report a not negligible decrease in the average quality of college graduates (around 6%) between 1960 and 2000 but substantially lower rates (around 1%) in the last decade of the last century.

The role of unemployment in relative labor supply determination. Unemployment could imply a bias in our SBTC estimates to the extent it affects the relative use of skilled and unskilled workers over time. OECD (2020) data on unemployment by education, for population between 25 and 64, show that changes from 1995 to 2005 are respectively 0.7, 0.8 and 0.25% for people with education levels below upper secondary, upper secondary and tertiary attainment. We cannot precisely attribute these unemployed to each sector (a share of them may be hired from different sectors) so that recomputing the labor ratios is not straightforward. For the sake of discussion, if we consider an homogeneous change to the labor supplies due to this unemployment (that means adding these percentage changes proportionally to each sector) a raw computation according to (9) shows that a negative additional change in skill workers ratio of 0.5% would imply an additional yearly difference of 0.025% in the estimated SBTC. This means that this channel would not dramatically change our results.

[19]We wish to thank David Rivers for insightful discussions on this point.

6 Economic inefficiency analysis

Having quantified the country-sector MRTS, as well as the contribution of SBTC to the MRTS evolution, independently of relative wages, we are able to study the country-sector discrepancy between (estimated) MRTS and (observed) relative wages. The 2005 level of such difference, as measured through the index $\ln \widehat{\tau}^{s,u}_{cj,T}$ in Equation (27), is reported in Tables 4 and 5. The dominance of negative values suggests that the wage ratio is usually larger than the MRTS. In the presence of decreasing MPs, this would call, at given wage levels, for an increase in the relative use of unskilled labor.[20]

The third panel of Figures 4 and 5 contrasts the country and sector averages of the MRTS change to the corresponding change in the wage ratio (as suggested by Equation (27)), plotted against the bisector line.[21] The skill premium grows less than the MRTS in most countries and sectors. It grows more, on average, in only a few countries such as the US, Germany, Korea, Ireland and Japan. Countries such as Romania, Malta, and Indonesia show extremely high MRTS growth, in sectors such as "Retail trade", "Hotels & restaurants", "Leather & footwear" and "Air transport". We observe decreasing MRTS in only a few sectors. This is the case of "Machinery", among manufacturing sectors, and "Business Services", "Utilities", among non-manufacturing sectors.

The first two panels in Figures 4 and 5 show that the estimated SBTC is positive and larger than the change in the wage ratio (above the bisector line) in all countries but India and in the vast majority of sectors. By contrast, the average FA component is negative in many countries and sectors, often growing less, or decreasing more, than the wage premium. We detect substantial heterogeneity across sectors and countries, with FA taking the lead in several cases as the driver of MRTS changes.

To provide an assessment of the country-sector evolution of the deviation from economic efficiency in the use of skilled labor, relative to unskilled labor, we compute the measure $\widehat{\Theta}^{s,u}_{cj,tT}$ defined in Equation (29). As mentioned, positive (negative) values refer to increasing (decreasing) discrepancy between relative wages and MRTS, and thus to increasing (decreasing) economic inefficiency in the relative use of skilled and unskilled labor. Under the hypothesis of perfectly competitive labor markets, or in the presence of distortions affecting the skilled and unskilled labor markets symmetrically, MRTS and wage ratio would coincide in levels and growth rates. However, this is not the case in our estimates.

The values reported in Tables 4 and 5 indicate an average 10.5% decrease in economic inefficiency, with considerable cross-country and cross-industry differences. The highest gains in efficiency are estimated for Malta, Romani and, notably, for Italy, followed by Cyprus, Estonia and Brazil. Inefficiency is instead found to grow substantially in India (45%). Interestingly, economic efficiency grows also in countries such Spain and Luxembourg and, to a lesser extent, Germany, Netherlands and Australia. At the sectoral level,

[20] Although investigating whether MP are increasing or decreasing is beyond the scope of this paper, Appendix C shows that, unlike capital, for which the decreasing MP hypothesis is broadly verified, increasing (either skilled or unskilled) labor is not associated with lower MP (neither in levels nor in growth rates).

[21] Appendix C highlights how wages and MPs do not match in the vast majority of cases. The country-sector wage level is often higher than the corresponding MP; this is even more true in the skilled labor market, with a slight tendency to balancing across the period. Moreover, the overall increase in MP in the skilled labor market (.041 on average) is associated with shrinking (PPP) real wages ($-.31$ on average) and growing used quantity (.53 on average). Interestingly, while the wages of unskilled workers' shrink at a similar rate, a slight overall decrease in their MP is detected ($-.01$ an average), associated with a slight growth in U (.02).

increases are reported only in the two non-manufacturing industries "Business Services" (around 10%) and "Utilities" (+0.4%). The largest inefficiency reductions are in the "Basic and fabricated metals" (around -25%) and "Agriculture" (-20%). Notably, efficiency improvements are higher in non-OECD countries and in manufacturing sectors.

The above evidence confirms the idea that SBTC estimates obtained through the indirect approach, under the hypothesis of perfectly competitive labor markets, are a composition of pure SBTC and economic inefficiency (discrepancy between MRTS and skill premium).

6.1 Inefficiency markers.

Previous analysis reveals increasing economic inefficiency in the use of skilled labor, relative to unskilled labor, in the 95-05 decade. While we find SBTC to be primarily responsible for the MRTS increase, the existence, and widening, of the gap between MRTS and relative wages arguably depends on sluggish adjustment in wages eventually related to differences (between the skilled and unskilled workers) in terms of bargaining power and institutional settings (i.e., minimum wage).

To give this dimension an order of magnitude, we carry out a simple econometric analysis to uncover potential correlation between change in economic inefficiency and labor market characteristics capturing the easiness of adjustment of wages.

To capture potential asymmetries in the attitude of skilled workers' wages to adjust to their MP, relative to unskilled workers, we use data from Visser (2019). In particular, we rely on the 1995-2005 averages of:

- *Minimum Wage Setting*: a variable reflecting "the (increasing) degree of government intervention and discretion in setting the minimum wage, or - reversely - the degree to which the government is bound in its decisions by unions and employers, and/or a fixed or pre-determined rule" (Visser, 2019). The variable is coded from 0 (i.e., No statutory minimum wage) to 9 (minimum wage set by government).

- *Union Density*: a variable measuring union membership as the proportion of wage and salary earners in total employment that are members of a trade union. This variable measures better than the raw number of union members the weight of the union since other population groups like retired people are often the biggest part of members.

Indeed, Di Nardo et al. (1996), Acemoglu et al., (2001) and Autor et al. (2016) find wage ratio compression effects associated with the action of trade unions and labor market institutions. In particular, Card et al. (2020) estimate that trade unions reduce wage inequality by 10%. As we have reported decreasing wage premia and increasing MRTS, we expect both variables to positively correlated with our measure of change in economic inefficiency, $\widehat{\Theta}^{s,u}_{cj,tT}$.

We also ask whether economic inefficiency at the country-sector level is affected by its initial values by regressing $\widehat{\Theta}^{s,u}_{cj,tT}$ against the absolute value of $\ln \widehat{\tau}^{s,u}_{cj,1995}$ (as the extent of economic inefficiency does not depend on its sign). Under a standard convergence interpretation, the discrepancy between MRTS and relative wages is expected to decrease in country-sectors characterized by higher initial values; visual inspection (see Figure 6) points to such a pattern.

We estimate the following equation as a cross-section (1995-2005 averages):

$$\widehat{\Theta}^{s,u}_{c,j} = \alpha + \beta_1 |\ln \widehat{\tau}^{s,u}_{cj,1995}| + \beta_2 \text{IT Exposure}_j * \text{Institutions}_c + \epsilon_{c,j}, \tag{30}$$

where *IT Exposure* stands for the US exposure to SBTC in sector j, as proxied by US IT employment change in sector j, computed on EUKLEMS (2019) data. *Institutions* refers to our measures of institutional bound *Minimum Wage Setting* and *Union Density*.

The logic behind this interaction is that we should expect that institutions should have a positive impact on this wedge (by compressing the wage ratios), that should be growing in sector-level more exposed to skill-biased technological change.

Table 8 summarizes the regression output. Given that the dependent variable is a generated regressor, we use boostrapped standard errors with 1000 replications. A first result is the strong evidence of convergence in economic inefficiency. According to Figure 6, economic inefficiency seems to decrease more in more inefficient country-sectors and slightly grow in less inefficient ones. This is confirmed by the econometric analysis and persists in all the specifications adopted.[22]

Turning to labor market characteristics, while we lose parts of the observations due to missing union data, the interacted variables seems to exert, as expected (Machin, 1997, Koeniger et al., 2007, Checchi and Garcia-Penalosa, 2008), a strongly significant positive effect suggesting the presence of labor market institutions to be more binding, in terms of wage and/or labor quantity adjustment, in sectors that are more exposed to technological advancement. Alternative controls, such as adjusted coverage of collective bargaining and extra-coverage (i.e., difference between density and coverage), yield similar results. Even if these are just correlations we may interpret this as an ex post validation of our gap measure between wages and productivities ratios. This strong positive correlation with binding institutions seems to go in the direction of a wage compressing role of the latter.

6.2 The wage inequality effect of restoring economic efficiency through wage adjustment

In this Section we carry out a counterfactual experiment in which we imagine wages to fully adjust to MPs, so as to re-establish economic efficiency. The measure in Equation (28) informs us about the change in the wage ratio that would have allowed the achievement of economic efficiency via relative wages. How would the wage distribution change under full adjustment of wages to MP? While, as seen, both the real wage and the amount of unskilled labor shrank on average, in the case of skilled labor,

[22] Given we are using a sample of countries and sectors allowing a larger degree of heterogeneity, with respect for instance simple OECD, potential issues of benchmarking bias as in Ciccone and Papaioannou (2016) should make us consider as our results are more conservative than possible alternatives, because in this case they found a downward bias.

we report declining real wages and increasing amounts used. This suggests the presence of asymmetric distortions in the labor market, which potentially result in less wage inequality with respect to a fully competitive market.

As we know the wages of high and low skilled workers' in 1995 and 2005, as well as the quantity of S and U in each country-sector, we can derive an overall wage distribution $\Omega_t^W = F\left(W^s_{cj,t}, \theta^s_{cj,t}; W^u_{cj,t}, \theta^u_{cj,t}\right)$, with $t = 1995, 2005$ and the θs denoting the share of skilled and unskilled labor in the given country-sector.

On this basis, we can compute the actual measures of wage inequality, by using sectoral wages weighted by relative PPPs to take into account differences among sectors and countries, and contrast them with a counterfactual situation in which wages are assumed to vary in the same proportion of MP in each country-sector, in both the skilled and the unskilled labor market, that is, in $\Delta \ln \widetilde{W}^s_{cj,tT} = \Delta \ln m^s_{cj,tT}$ and $\Delta \ln \widetilde{W}^u_{cj,tT} = \Delta \ln m^u_{cj,tT}$. Here the amount of S and U is fixed, while wages completely follow marginal products.

Figure 7 displays the resulting actual (solid line) and counterfactual (dashed line) 2005 kernel densities. Unsurprisingly, more differences are in the left tail, entailing that lower wages would have been even lower if they followed MP. This is consistent with the remark by Caselli (2016) concerning the potential underestimation of SBTC through the indirect approach when labor market institutions are relatively more egalitarian, as in the case of more developed countries, compared to emerging ones.

To synthesize the effect on wage inequality, we can use concentration measures such as the Gini index, the Atkinson or the decile and interquartile range. These are reported in Table 9. Computations use labor stocks as weights. In this way, each sector represents its relative percentage in the economy. Being expressed in PPPs, wages are comparable across countries.

The true Gini index computed on 2005 data amounts to 0.78. This fits quite well with the findings by Bourguignon and Morrisson (2002) and Milanovic (2012), who report world values around 0.70. While, consistently with Milanovic (2012), we report a declining wage concentration over the 95-05 decade, all measures point to increasing concentration under the counterfactual hypothesis that wages evolve proportionally to MP in both the skilled and unskilled labor market. The Gini index grows to 0.86, which is 10% points more than in 1995 (if we take the average of within country inequality, the implied change is bigger than 20%).

7 Conclusions

Despite widespread interest in the productivity consequences of the IT revolution, economic literature has so far produced very limited evidence on the magnitude of SBTC. Indeed, apart from the country-level contribution by Caselli (2016), reporting a relatively high SBTC for high-income countries, compared to low-income ones, no evidence at all has been made available at the country-sector level. Moreover, extant country-level estimates are inferred from the (observed) wage gap between skilled and unskilled workers (wage-based approach). However, as known, SBTC requires adjustment in terms of the skill premium and/or the relative use of skilled and unskilled labor (i.e., the labor ratio). As long as labor market frictions/ imperfections affect the skilled and unskilled labor markets asymmetrically, thereby, preventing adjustments: i) the wage structure is no longer informative of the ongoing technological progress and cannot be used to infer SBTC; technological progress can be, somehow paradoxically, associated with a less efficient use of labor.

In this paper, we ask to what extent has the recent wave of technological progress come in a "skill biased" form, and consequent to this, with asymmetric effects on the relative demand of skills and relative wages (i.e., skill premium). By doing it we try to relax two implicit assumptions behind the traditional estimation that are (i) a functional form for the aggregated production function (typically CES), and (ii) the assumption that relative wages capture marginal products.

These issues cannot be addressed using the indirect, wage-based, approach, which assumes the economic efficiency condition (equality between MRTS and relative wage) to be satisfied. Such an approach attributes the full deviation from the efficiency condition to technological progress, under the hypothesis that both the skilled and unskilled labor markets feature perfect competition, thereby neglecting the action of labor market frictions/distortions. To overcome this limitation, we presented an alternative approach that allows us to: i) obtain country-sector SBTC estimates that are net of the effect of FA; ii) disentangle the SBTC and FA contribution to the estimated change in MRTS; iii) quantify the discrepancy between the wage ratio (of skilled to unskilled workers) and the MRTS (i.e., economic inefficiency).

Estimation relies on nonparametric methods allowing for country-sector-year specific estimates of the MRTS between skilled and unskilled labor directly from country-sector-year data, which would be a daunting task with standard parametric techniques. This methodology allows us to depart from extant literature in retrieving SBTC directly from the production function, which in turn enables us to avoid conflating "true" SBTC and labor market distortions.

Based on WIOD data, our empirical analysis reveals that the MRTS between skilled and unskilled labor has been growing by 11.5% on average over the decade 1995-2005. By decomposing (through counterfactual analysis) the MRTS variation into a SBTC component and a FA component, we discover that most of the MRTS change is associated with SBTC, whose average effect on MRTS we quantify to be 13.7%. This value complements with the -2.2% that we find for FA to obtain the total change in the productivity ratio.

Interestingly, our average values disguise substantial heterogeneity across countries and sectors, with the FA effect sometimes dominating the SBTC component. Notably, this is the case in the US, for which we estimate a -20% decrease in MRTS associated with a conspicuously negative FA contribution (-49 %) and a positive SBTC (+29 %).

We also report a 10.5% overall decrease in economic inefificiency (i.e., discrepancy between MRTS and wage premium), although with substantial heterogeneity across sectors and countries. Interestingly, economic efficiency grows in all sectors but utilities and most OECD and non-OECD countries, with notable exceptions such as Germany, Netherlands, Australia, Canada. This suggests that SBTC can induce, in some cases, a less efficient use of labor (with respect to an ideal, economically efficient, situation in which the MRTS between skilled and unskilled workers equals the wage ratio) as long as it does not come with sufficiently higher skill premia and subsequent adjustment in the adopted quantity of skilled and unskilled labor.

Econometric analysis suggests that labor market institutions are more binding, in terms of wage and/or labor quantity adjustment, in sectors that are more exposed to technological advancement.

We finally used counterfactual analysis to show that full adjustment of wages to MPs would bring about an increase in wage inequality (0.81 against 0.76 in terms of the Gini index) mostly through further decreases in low wages.

In terms of policy implications, our work highlights the importance of wage adjustment when technological progress is skill biased, as the wave of technological progress occurred in the 1995-2005 decade. Indeed, by changing the relative productivity of skilled and unskilled workers, SBTC requires substantial adjustment in terms of both wage premia and quantity of skilled labor used, relative to unskilled labor. This questions the changing role of labor market institutions in the face of SBTC.

References

[1] Acemoglu, D. (2002). Directed technical change. *Review of Economic Studies*, 69, 781-810.

[2] Acemoglu, D. & Aghion. P. & Violante, G.L. (2001). Deunionization, technical change and inequality, *Carnegie-Rochester Conference Series on Public Policy*, 55, 229-264.

[3] Acemoglu, D. & Autor, D.H. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. *Handbook of Labor Economics*, Volume 4b, 1043-1171.

[4] Alesina, A. & Battisti, M. & Zeira, J. (2018). Technology and labor regulations: theory and evidence, *Journal of Economic Growth*, 23(1), 41-78.

[5] Autor, D. H., Katz, L. F., & Krueger, A. B. (1998). Computing inequality: have computers changed the labor market? *The Quarterly journal of economics*, 113(4), 1169-1213.

[6] Autor, DH. & Levy, F. & Murnane, R.J. (2003). The Skill-Content of Recent Technological Change: An Empirical Investigation, *Quarterly Journal of Economics*, 118, 1279-1333.

[7] Autor, D.H. & Manning, A. & Smith, L.C. (2016). The Contribution of the Minimum Wage to US Wage Inequality over Three Decades: A Reassessment, *American Economic Journal: Applied Economics*, 8 (1), 58-99.

[8] Bartelsman, E. & Haltiwanger, J. & Scarpetta, S. (2013), Cross-Country Differences in Productivity: The Role of Allocation and Selection, *American Economic Review*, 103(1), 305-334.

[9] Baqaee, D.R. & Farhi, E. (2020). Productivity and misallocation in general equilibrium, *Quarterly Journal of Economics*, 135(1), 105-163.

[10] Battisti, M. & Del Gatto, M. & Parmeter, C.F. (2018). Labor productivity growth: disentangling technology and capital accumulation, *Journal of Economic Growth*, 23(1), 111-143.

[11] Blankenau, W.F. & Cassou, S.P. (2011). Industry estimates of the elasticity of substitution and the rate of biased technological change between skilled and unskilled labour, *Applied Economics*, 43(23), 3129-3142.

[12] Bound, J. & Johnson, G. (1992). Changes in the Structure of Wages in the 1980's: An Evaluation of Alternative Explanations, *American Economic Review*, 82(3), 371-392.

[13] Bourguignon F. & Morrisson, C. (2002). Inequality among World Citizens: 1820-1992, *American Economic Review*, 92(4), 727-744.

[14] Brynjolfsson, E. (1993). The Productivity Paradox of Information Technology. *Communications of the ACM*, 36 (12), 66-77.

[15] Brynjolfsson, E., Rock, D., & Syverson, C. (2019). Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics. In *Economics of Artificial Intelligence: An Agenda, 23*, University of Chicago Press.

[16] Brynjolfsson, E., Rock, D., & Syverson, C. (2021). The productivity J-curve: How intangibles complement general purpose technologies. *American Economic Journal: Macroeconomics*, 13(1), 333-72.

[17] Buera, F. & Kaboski, J. & Rogerson, R. & Vizcaino, J. (2021). Skill-Biased Structural Change, *Review of Economic Studies*, forthcoming.

[18] Calligaris, S. & Del Gatto M. & Hassan, F. & Ottaviano, G. & Schivardi, F. (2018). The productivity puzzle and misallocation: an Italian perspective, *Economic Policy*, 33(96), 635?684.

[19] Card, D. (1992). The Effect of Unions on the Distribution of Wages: Redistribution or Relabelling? *NBER Working Papers*, 4195.

[20] Card, D. & Di Nardo, J. (2002). Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles, *Journal of Labor Economics*, 20(4), 733-783.

[21] Card, D. & Lemieux, T. (2001). Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis, *Quarterly Journal of Economics*, 116(2), 705-746.

[22] Card, D., Lemieux, T., & Riddell, W. C. (2020). Unions and wage inequality: The roles of gender, skill and public sector employment. *Canadian Journal of Economics*, 53(1), 140-173.

[23] Carneiro, P. & Lee, S. (2011). Trends in Quality-Adjusted Skill Premia in the United States, 1960-2000, *American Economic Review*, 101, 2309-2349.

[24] Caselli, F. (1999). Technological Revolutions, *American Economic Review*, 89(1), 78-102.

[25] Caselli, F. (2005). Accounting for Cross-Country Income Differences, in S.N. Durlauf and P. Aghion (Eds.) *Handbook of Economic Growth*, Elsevier.

[26] Caselli, F. (2016). *Technology Differences among Space and Time*, Princeton University Press.

[27] Caselli, F., & Coleman, W. J. (2006). The world technology frontier, *American Economic Review*, 96(3), 499-522.

[28] Checchi, D. & Garcia-Penalosa, C. (2008). Labour market institutions and income inequality, *Economic Policy*, 23(56), 602-649.

[29] Ciccone, A. & Papaioannou, E. (2016). Estimating Cross-Industry Cross-Country Interaction Models Using Benchmark Industry Characteristics, *NBER Working Papers*, No. 22368.

[30] Di Nardo, J. & Fortin, N.& Lemieux, T., (1996). Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach. *Econometrica*, 64, 1001- 1044.

[31] Dorazelski, U. & Jaumandreu, J. (2018). Measuring the Bias of Technological Change, *Journal of Political Economy*, 126(3), 1027-1084.

[32] Erumbam, A.A. & Gouma, R. & De Vries, G. &De Vries, K. & Timmer, M. (2012). *WIOD Socio-Economic Accounts (SEA): Sources and methods.* Groningen, April.

[33] Feenstra, R.C. & Inklaar, R.C., & Timmer, M. P. (2015). The next generation of the Penn World Table, *American Economic Review*, 105(10), 3150-3182.

[34] Foster McGregor, N. & Poschl, J. (2009). The Importance of Labour Mobility for Spillovers across Industries, Wivi Working papers no. 58.

[35] Freeman, R.B (1993). Labor Markets and Institutions in Economic Development, *American Economic Review*, 83(2), 403-408.

[36] Freeman, R.B. & Katz, L.F. (1995). *Differences and Changes in Wage Structures. Chicago*, University of Chicago Press.

[37] Goos, M., & Manning, A. (2007). Lousy and lovely jobs: The rising polarization of work in Britain, *Review of Economics and Statistics*, 89(1), 118-133.

[38] Goos, M., & Manning, A. & Salomons, A. (2014). Explaining job polarization: Routine-biased technological change and offshoring, *American Economic Review* (104), 8, 2509-2526.

[39] Henderson, D.J. (2009). A Non-parametric Examination of Capital-Skill Complementarity, *Oxford Bulletin of Economics and Statistics*, 71(4), 519-538.

[40] Henderson, D. J., & Parmeter, C.F. (2015). *Applied Nonparametric Econometrics.* Cambridge University Press.

[41] Hsieh, C.T. & Klenow, P.J. (2009). Misallocation and Manufacturing TFP in China and India, *Quarterly Journal of Economics*, 124(4), 1403-1448.

[42] Juhn, C. & Murphy, K.M. & Pierce, B. (1992). Wage Inequality and the Rise in Returns to Skill, *Journal of Political Economy*, 101(3), 410-442.

[43] Kambourov, G. & Manovskii, I. (2008). Rising Occupational and Industry Mobility in the United States: 1968-97, *International Economic Review*, 49(1), 41-79.

[44] Karabarbounis, L., & Neiman, B. (2014). The global decline of the labor share, *Quarterly Journal of Economics*, 129(1), 61-103.

[45] Katz, L. F. & Murphy, K. M. (1992). Changes in relative wages, 19631987: Supply and demand factors, *Quarterly Journal of Economics*, 107, 35-78.

[46] Koeniger, W. & Leonardi, M. (2007). Capital deepening and wage differentials: Germany versus US, *Economic Policy*, 22(49), 72-116.

[47] Koeniger, W. & Leonardi, M. & Nunziata, L. (2007). Labor Market Institutions and Wage Inequality *Industrial Labor Relations Review*, 60(3), 340-356.

[48] Krusell, P. & L. Ohanian & J.V. Ros-Rull & Violante, G.L. (2000). Capital Skill Complementarity and Inequality: A Macroeconomic Analysis, *Econometrica*, 68, 1029-1054.

[49] Lee, D.S. (1999). Wage Inequality in the United States During the 1980s: Rising Dispersion or Falling Minimum Wage? *Quarterly Journal of Economics*, 114(3), 977-1023.

[50] Lemieux, T. (2008). The changing nature of wage inequality, *Journal of Population Economics*, 21, 21-48.

[51] Levy, F. & Murname, R. (1992). U.S. Earnings Levels and Earnings Inequality: A Review of Recent Trends and Proposed Explanations, *Journal of Economic Literature*, 30(3), 1333-1381.

[52] Li, C. & Ouyang, D. & Racine, J.S. (2009). Nonparametric regression with weakly dependent data: The discrete and continuous regressor case, *Journal of Nonparametric Statistics*, 21, 697-711.

[53] Li, Q. & Racine, J.S. (2004). Cross-validated local linear nonparametric regression, *Statistica Sinica*, 14, 485-512.

[54] Li, Q., & Racine, J.S. (2007). *Nonparametric Econometrics*. Princeton University Press.

[55] Loader, C. (1999). *Local Regression and Likelihood*, Springer-Verlag New York.

[56] Machin, S. (1997). The decline of labour market institutions and the rise in wage inequality in Britain, *European Economic Review*, 41(3-5), 647-657.

[57] Milanovic, B. (2012). Global Inequality Recalculated and Updated: The Effect of New PPP Estimates on Global Inequality and 2005 Estimates, *Journal of Economic Inequality*, 10, 1-18.

[58] Neffke, F.M.H. & Otto, A. & Weyh, A. (2017). Inter-industry labor flows, *Journal of Economic Behavior and Organization* 142, 275-292.

[59] OECD (2018). Decoupling of Wages from Productivity: what Implications for Public Policies? In *OECD Economic Outlook*, Vol. 2018, Issue 1, Chapter 2, OECD Publishing, Paris.

[60] OECD (2020). Education at a glance: Educational attainment and labour-force status.

[61] Park, J. & Wood, I.B & Jing, E. & Nematzadeh, A. & Ghosh, S. & Conover, M.D. & Ahn, Y.-Y. (2019). Global labor flow network reveals the hierarchical organization and dynamics of geo-industrial clusters, *Nature Communications*, 10, 3449.

[62] Restuccia, D. & R. Rogerson (2008). Policy distortions and aggregate productivity with heterogeneous establishments, *Review of Economic Dynamics*, 11(4), 707-720.

[63] Restuccia, D. & Rogerson, R. (2013). Misallocation and Productivity, *Review of Economic Dynamics*, 16, 1-10.

[64] Robinson, P.M. (1983). Nonparametric estimators for time series, *Journal of Time Series Analysis*, 4, 185-207.

[65] Rossi F. (2019). The Relative Efficiency of Skilled Labor Across Countries: Measurement and Interpretation, unpublished, Johns Hopkins University.

[66] Solow, R. (1987). We'd Better Watch Out. *New York Times Book Review.*

[67] Su, L. & Ullah, A. (2008). Local polynomial estimation of nonparametric simultaneous equations models, *Journal of Econometrics*, 144(1), 193-218.

[68] Timmer , M.P. (2012). The World Input-Output Database (WIOD): Contents, Sources and Methods, WIOD, Working Paper no. 10.

[69] Timmer, M.P. & Dietzenbacher, E. & Los, B. & Stehrer, R. & de Vries, G.J. (2015). An Illustrated User Guide to the World Input-Output Database: the Case of Global Automotive Production, *Review of International Economics*, 23, 575-605.

[70] Violante, G.L. (2016). Skilled-biased Technical Change, in *The New Palgrave Dictionary of Economics*, ed. by Durlauf, S.N. and Blume, L.E. New York: Palgrave MacMillan.

[71] Visser, J. (2019). *ICTWSS Data base. version 6.1.*, Amsterdam Institute for Advanced Labour Studies AIAS, Amsterdam.

[72] Zeira, J. (1998). Workers, Machines, and Economic Growth, *Quarterly Journal of Economics*, 113(4), 1091-1117.

Table 1: Descriptive statistics by country (logarithmic values)

		Y		K		S		U		S/U gr.	
COUNTRY	Code	1995	2005	1995	2005	1995	2005	1995	2005	95-05	#obs
Australia	AUS	8.74	8.98	9.41	10.01	2.90	3.30	5.38	5.38	41%	27
Austria	AUT	8.38	8.60	9.11	9.27	2.07	2.70	4.60	4.54	69%	25
Belgium	BEL	8.46	8.59	9.22	9.46	2.11	2.25	4.41	4.32	26%	23
Bulgaria	BGR	5.84	6.08	6.59	6.99	1.17	1.42	4.45	4.31	41%	29
Brazil	BRA	9.61	9.85	10.35	10.85	4.70	5.13	7.31	7.49	28%	28
Canada	CAN	8.98	9.27	9.51	9.73	3.43	3.81	5.58	5.62	34%	26
Cyprus	CYP	4.86	5.06	5.50	5.60	0.62	0.47	1.90	1.92	-19%	24
Czech Republic	CZE	7.31	7.40	8.34	8.67	2.40	2.46	5.10	4.90	26%	29
Germany	DEU	10.49	10.55	11.09	11.19	4.90	4.92	6.54	6.34	22%	27
Denmark	DNK	7.81	7.90	8.39	8.66	2.17	2.47	4.02	3.84	49%	26
Spain	ESP	9.28	9.55	9.96	10.41	3.99	4.72	5.74	5.82	67%	26
Estonia	EST	4.55	5.34	5.28	6.29	1.68	1.60	2.79	2.67	10%	28
Finland	FIN	7.53	7.83	8.22	8.38	2.50	2.67	3.76	3.77	16%	25
France	FRA	9.83	10.09	10.37	10.55	4.24	4.52	6.03	5.84	48%	27
United Kingdom	GBR	9.71	9.86	10.02	10.26	4.44	4.61	6.24	5.93	49%	24
Greece	GRC	7.48	7.80	7.98	8.35	2.28	2.62	4.54	4.54	35%	23
Hungary	HUN	6.76	7.01	8.02	8.25	2.37	2.60	4.72	4.66	28%	29
Indonesia	IDN	9.17	9.44	9.94	10.37	3.45	4.45	6.87	7.37	46%	20
India	IND	9.43	10.06	10.33	11.07	5.83	6.57	8.37	8.58	50%	27
Ireland	IRL	6.81	7.20	7.29	7.86	1.63	2.43	3.49	3.56	76%	21
Italy	ITA	10.04	10.10	10.87	11.14	3.47	3.98	6.66	6.65	54%	24
Japan	JPN	11.15	11.07	11.83	12.08	5.67	5.71	7.55	7.25	31%	22
Korea, Republic of	KOR	9.45	9.80	10.26	10.60	5.23	5.68	6.50	6.25	70%	28
Lithuania	LTU	6.24	6.95	6.94	7.79	1.92	2.00	3.26	3.27	10%	29
Luxembourg	LUX	5.45	5.91	6.09	6.65	-0.45	0.02	1.49	1.77	23%	20
Latvia	LVA	4.10	4.89	4.90	5.62	1.03	1.14	2.55	2.77	1%	30
Mexico	MEX	8.91	9.24	9.57	9.73	4.74	5.00	6.71	6.97	-1%	28
Malta	MLT	4.35	4.50	5.04	5.35	-1.65	-1.29	1.46	1.46	29%	24
Netherlands	NLD	8.66	8.85	9.17	9.35	2.72	3.21	4.83	4.67	63%	24
Poland	POL	8.09	8.57	8.78	9.28	3.47	3.91	5.95	5.78	57%	29
Portugal	PRT	7.62	7.87	8.36	8.81	1.47	1.85	4.68	4.69	37%	23
Romania	ROU	7.01	7.19	7.67	8.06	2.19	2.54	5.50	5.45	39%	28
Russia	RUS	9.47	9.75	10.30	10.42	4.65	4.72	7.44	7.28	24%	25
Slovak Republic	SVK	6.35	6.72	7.04	7.62	1.59	1.65	4.15	4.00	23%	28
Slovenia	SVN	6.06	6.30	6.85	7.45	0.96	1.15	3.24	3.03	40%	29
Sweden	SWE	8.13	8.39	8.65	8.99	2.18	2.63	4.56	4.40	62%	24
Turkey	TUR	8.87	9.48	9.55	10.08	3.07	3.78	6.15	6.33	49%	25
United States	USA	11.48	11.72	11.93	12.14	6.15	6.21	7.75	7.62	21%	21
Avg/Tot		7.96	8.26	8.65	9.04	2.82	3.15	5.06	5.03	36%	975

Table 2: Descriptive statistics by sector (logarithmic values)

		Y		K		S		U		S/U gr.	
SECTOR	Short	1995	2005	1995	2005	1995	2005	1995	2005	95-05	#obs
Food, beverage & tobacco	FD	8.67	8.80	9.46	9.71	3.32	3.57	5.69	5.57	40%	37
Textile products	TX	8.23	8.17	8.76	8.79	3.05	3.06	5.56	5.15	48%	37
Leather & footwear	LF	6.41	6.12	6.95	6.89	1.35	1.15	3.81	3.21	45%	33
Wood products	WO	7.02	7.34	7.55	7.98	2.07	2.41	4.48	4.44	43%	38
Paper, printing & publishing	PP	7.90	8.18	8.53	8.91	2.71	3.01	4.94	4.88	39%	37
Coke & refined petroleum	CP	6.30	6.22	7.48	7.85	1.02	1.11	3.12	2.90	32%	22
Chemical products	CH	7.97	8.33	8.69	9.02	2.67	2.81	4.75	4.59	32%	29
Rubber & plastics	RP	7.15	7.68	7.75	8.25	2.06	2.47	4.34	4.41	40%	38
Non-metallic mineral products	NM	7.83	8.20	8.70	8.96	2.43	2.63	4.72	4.56	38%	36
Basic & fabricated metal	BM	8.37	8.71	9.16	9.40	3.11	3.49	5.41	5.43	38%	38
Machinery	MA	7.87	8.29	8.24	8.56	2.82	3.02	5.07	4.88	44%	37
Electrical & optical equipment	EL	7.80	8.42	8.21	8.72	2.65	2.98	4.90	4.88	40%	33
Transport Equipment	TR	7.64	8.12	8.17	8.64	2.45	2.82	4.69	4.68	41%	37
Other manufacturing	OT	7.29	7.64	7.59	8.03	2.39	2.77	4.67	4.68	39%	35
Motor vehicle & fuel trade	mv	7.78	8.16	8.23	8.63	2.89	3.37	5.09	5.27	24%	34
Wholesale trade	wt	9.08	9.50	9.34	9.77	3.94	4.34	6.12	6.20	33%	33
Retail trade	rt	9.02	9.40	9.34	9.79	4.42	4.88	6.66	6.80	37%	37
Land transport	lt	8.62	8.87	9.94	10.32	3.72	4.05	5.93	5.95	32%	38
Water transport	wa	5.99	6.13	7.33	7.62	0.92	1.04	3.06	2.88	35%	32
Air transport	at	5.97	6.31	7.30	7.63	1.02	1.39	3.19	3.24	42%	33
Transport services	ts	7.91	8.31	9.18	9.72	2.50	3.05	4.66	4.92	34%	37
Post & telecommunications	pt	8.10	8.84	8.92	9.69	2.79	3.35	5.08	5.18	49%	32
Real estate	re	9.55	10.10	11.83	12.19	4.96	5.51	5.59	5.99	23%	4
Business services	bs	8.94	9.47	9.07	10.10	4.97	5.61	5.61	6.03	19%	33
Agriculture, forestry & fishing	ag	9.52	9.64	10.45	10.73	4.12	4.34	7.39	7.07	54%	32
Mining & quarrying	mq	7.10	7.18	8.40	8.55	1.81	1.84	4.18	3.92	33%	31
Utilities	ut	8.03	8.20	9.80	10.21	2.90	3.26	5.02	4.81	55%	20
Construction	co	8.92	9.18	8.98	9.44	3.80	4.11	6.35	6.50	19%	30
Hotels & restaurants	hr	8.36	8.63	8.69	9.27	3.27	3.79	6.00	6.22	35%	38
Financial services	fs	9.07	9.38	9.26	9.66	4.21	4.64	5.22	5.16	50%	24
Avg/Tot		7.95	8.25	8.71	9.10	2.88	3.20	5.04	5.01	36%	975

Table 3: Estimated MRTS (average values).

	P25	Mean	P50	P75
$\widehat{MRTS}^{s,u}_{cj,t}$	1.101	1.487	1.356	1.677
	(0.275)	(0.595)	(0.368)	(0.545)
$\widehat{MRTS}^{s,u}_{cj,T}$	1.204	1.819	1.442	1.761
	(0.241)	(0.716)	(0.345)	(0.551)
$\widetilde{MRTS}^{s,u}_{cj,T}$	1.267	1.815	1.530	1.847
	(0.062)	(0.384)	(0.120)	(0.218)

[a] Standard errors in parentheses. P=%ile.
[b] Computation of Equation (21) for initial, final and counterfactual periods.

Table 4: Decomposition by country.

COUNTRY	$\Delta \ln \widehat{MRTS}^{s,u}_{cj,tT}$	$\Delta \ln \widetilde{\mathbb{B}}^{s,u}_{cj,tT}$	$\Delta \ln \widetilde{\mathbb{F}}^{s,u}_{cj,tT}$	$\Delta \ln \mathbb{W}^{s,u}_{cj,tT}$	$\widehat{\Theta}^{s,u}_{cj,tT}$	$\ln \widehat{\tau}^{s,u}_{cj,T}$	# obs
AUS	0.044	0.126	-0.082	0.055	0.032	-0.099	27
AUT	0.101	0.126	-0.025	-0.034	-0.099	-0.308	25
BEL	0.086	0.122	-0.036	-0.034	-0.063	-0.304	23
BGR	0.137	0.162	-0.025	-0.095	-0.179	-0.533	27
BRA	0.104	0.082	0.022	-0.104	-0.262	-0.803	20
CAN	0.109	0.107	0.002	0.074	0.050	0.247	23
CYP	0.310	0.311	-0.002	-0.055	-0.348	-0.717	22
CZE	0.086	0.113	-0.027	0.007	-0.039	-0.300	29
DEU	-0.068	0.023	-0.091	0.019	0.006	-0.145	26
DNK	0.023	0.095	-0.072	0.033	0.032	-0.038	25
ESP	0.174	0.082	0.093	-0.065	0.165	-0.051	24
EST	0.193	0.319	-0.127	-0.112	-0.271	-0.694	25
FIN	0.112	0.118	-0.007	0.111	-0.022	-0.155	25
FRA	-0.019	0.074	-0.093	-0.176	-0.101	-0.191	27
GBR	-0.008	0.063	-0.070	-0.042	-0.007	0.042	24
GRC	0.019	0.106	-0.087	-0.026	-0.013	-0.094	23
HUN	0.127	0.115	0.013	-0.026	-0.144	-0.610	29
IDN	0.454	0.148	0.307	0.032	-0.115	-1.112	17
IND	0.452	-0.062	0.514	0.087	0.452	0.154	13
IRL	-0.034	0.165	-0.199	0.111	-0.010	-0.060	21
ITA	0.352	0.148	0.205	-0.195	-0.381	-0.633	23
JPN	-0.179	0.130	-0.309	-0.012	-0.142	0.288	19
KOR	-0.246	0.056	-0.302	0.037	-0.043	0.177	23
LTU	0.152	0.233	-0.082	-0.098	-0.226	-0.608	29
LUX	0.240	0.281	-0.040	0.358	0.127	-0.168	19
LVA	0.177	0.198	-0.020	-0.100	-0.247	-0.674	25
MEX	0.218	0.079	0.139	-0.081	0.091	-0.184	23
MLT	0.508	0.284	0.224	-0.102	-0.610	-1.630	22
NLD	0.052	0.100	-0.049	0.078	0.024	-0.090	24
POL	0.029	0.099	-0.071	-0.036	0.042	-0.113	28
PRT	0.105	0.088	0.018	-0.102	-0.206	-0.759	22
ROU	0.295	0.097	0.198	-0.108	-0.400	-0.657	27
RUS	0.096	0.138	-0.041	-0.021	-0.051	-0.261	21
SVK	0.057	0.103	-0.046	0.076	-0.021	-0.282	28
SVN	0.065	0.139	-0.074	-0.025	-0.090	-0.542	27
SWE	0.066	0.123	-0.057	-0.156	-0.073	-0.211	24
TUR	0.130	0.117	0.013	-0.053	-0.183	-0.631	24
USA	-0.195	0.290	-0.486	0.046	-0.117	0.332	19
NON OECD	0.163	0.150	0.013	-0.044	-0.149	-0.520	435
OECD	0.057	0.120	-0.062	-0.005	-0.048	-0.159	467
Avg/Tot	0.115	0.137	-0.021	-0.024	-0.105	-0.370	902

$\Delta \ln \widehat{MRTS}^{s,u}_{cj,tT}$ = Estimated Marginal Rate of Technical Substitution (growth rate); $\Delta \ln \widetilde{\mathbb{B}}^{s,u}_{cj,tT}$ = Counterfactual Skill Biased Technical Change Effect (growth rate); $\Delta \ln \widetilde{\mathbb{F}}^{s,u}_{cj,tT}$ = Counterfactual Factor Accumulation Effect (growth rate): see decomposition in Equation (23). $\Delta \ln \mathbb{W}^{s,u}_{cj,tT}$ = Wage premium growth rate; $\widehat{\Theta}^{s,u}_{cj,tT}$ = change in economic efficiency (decreasing values denoting increasing efficiency): see Equation (29); $\ln \widehat{\tau}^{s,u}_{cj,T}$ = Economic inefficiency in 2005: see Equation (27). Unweighted averages reported.

Table 5: Decomposition by sector.

COUNTRY	$\Delta \ln \widehat{MRTS}^{s,u}_{cj,tT}$	$\Delta \ln \widetilde{\mathbb{B}}^{s,u}_{cj,tT}$	$\Delta \ln \widetilde{\mathbb{F}}^{s,u}_{cj,tT}$	$\Delta \ln \mathbb{W}^{s,u}_{cj,tT}$	$\widehat{\Theta}^{s,u}_{cj,tT}$	$\ln \widehat{\tau}^{s,u}_{cj,T}$	# obs
Agriculture	0.217	0.079	0.138	-0.056	-0.202	-0.393	24
Air transport	0.225	0.219	0.006	0.067	-0.156	-0.519	31
Basic and fabricated metal	0.192	0.256	-0.064	-0.050	-0.249	-0.403	37
Business service	-0.085	0.047	-0.132	0.017	0.101	-0.347	29
Chemical product	0.024	0.100	-0.077	-0.045	-0.074	-0.414	27
Coke and refined products	0.102	0.144	-0.041	-0.025	-0.127	-0.544	18
Construction	0.238	0.112	0.126	-0.026	-0.032	-0.199	24
Electrical and optical	0.018	0.095	-0.077	-0.024	-0.140	-0.250	31
Financial services	-0.168	0.088	-0.256	-0.016	-0.055	-0.065	24
Food, beverage, tobacco	0.118	0.128	-0.010	-0.053	-0.104	-0.314	35
Hotels and restaurants	0.360	0.153	0.207	-0.084	-0.048	-0.314	31
Land transport	0.180	0.209	-0.029	0.010	-0.002	-0.214	35
Leather and footwear	0.164	0.174	-0.010	-0.073	-0.169	-0.522	32
Machinery	-0.030	0.083	-0.113	-0.024	-0.101	-0.152	35
Mining and quarrying	0.067	0.126	-0.060	-0.057	-0.076	-0.490	29
Motor vehicle	0.138	0.079	0.059	-0.011	-0.074	-0.215	34
Non-metallic min	0.026	0.129	-0.104	-0.033	-0.057	-0.363	35
Other manufacturing	0.091	0.126	-0.035	-0.030	-0.123	-0.247	33
Paper, printing	0.040	0.150	-0.109	-0.046	-0.123	-0.332	36
Post and telecommunication	0.123	0.156	-0.033	0.087	-0.038	-0.286	32
Real estate	0.408	0.310	0.098	-0.040	-0.448	-1.415	2
Retail trade	0.245	0.036	0.209	-0.065	-0.034	-0.001	28
Rubber and plastic	0.122	0.108	0.013	-0.042	-0.184	-0.376	38
Textile products	0.052	0.120	-0.069	-0.079	-0.186	-0.335	35
Transport Equipment	0.128	0.149	-0.020	-0.005	-0.062	-0.282	35
Transport services	0.108	0.135	-0.028	0.045	-0.066	-0.424	37
Utilities	-0.053	0.081	-0.133	-0.032	0.004	-0.475	19
Water transport	0.143	0.179	-0.036	0.058	-0.082	-0.587	29
Wholesale trade	0.192	0.122	0.070	-0.059	-0.154	-0.320	30
Wood products	0.075	0.204	-0.128	-0.045	-0.077	-0.299	37
Non-Manufacturing	0.137	0.126	0.011	-0.005	-0.063	-0.329	438
Manufacturing	0.081	0.142	-0.061	-0.041	-0.128	-0.337	464
Avg/Tot	0.115	0.137	-0.021	-0.024	-0.105	-0.370	902

$\Delta \ln \widehat{MRTS}^{s,u}_{cj,tT}$ = Estimated Marginal Rate of Technical Substitution (growth rate); $\Delta \ln \widetilde{\mathbb{B}}^{s,u}_{cj,tT}$ = Counterfactual Skill Biased Technical Change Effect (growth rate); $\Delta \ln \widetilde{\mathbb{F}}^{s,u}_{cj,tT}$ = Counterfactual Factor Accumulation Effect (growth rate): see decomposition in Equation (23). $\Delta \ln \mathbb{W}^{s,u}_{cj,tT}$ = Wage premium growth rate; $\widehat{\Theta}^{s,u}_{cj,tT}$ = change in economic efficiency (decreasing values denoting increasing efficiency): see Equation (29); $\ln \widehat{\tau}^{s,u}_{cj,T}$ = Economic inefficiency in 2005: see Equation (27). Unweighted averages reported.

Table 6: SBTC: broad validation

	(1)	(2)	(3)	(4)	(5)
Specialization * IT share (World)	0.385^{***}	0.402^{***}			
	(0.096)	(0.103)			
Specialization * IT share (US)			0.164^{***}		
			(0.058)		
Specialization * Robot stock (World)				0.135^{*}	
				(0.082)	
Concentration * IT share (World)					0.419^{**}
					(0.178)
Constant	0.108^{***}	0.115^{***}	0.121^{***}	0.126^{***}	0.021
	(0.008)	(0.016)	(0.016)	(0.016)	(0.048)
Sector controls	N	Y	Y	Y	Y
N	902	902	902	527	902
R^2	0.08	0.10	0.08	0.02	0.03

Estimation of Equation (30). Dep. Variable: $\Delta \ln \widetilde{\mathbb{B}}^{s,u}_{cj,tT}$ = Counterfactual Skill Biased Technical Change Effect (see Equation (24). Specialization and Concentration are country-specific variables computed on a output basis. IT share is a sector-specific variable computed on EUKLEMS IT capital. Robot stock is a sector-specific variable computed on IFR Robot stock data. Bootstrapped standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Robustness checks and alternative specifications (yearly changes)

	$\Delta\ln\widehat{MRTS}^{s,u}_{cj,tT}$	$\Delta\ln\widetilde{\mathbb{B}}^{s,u}_{cj,tT}$	$\Delta\ln\widetilde{\mathbb{F}}^{s,u}_{cj,tT}$	$ln\mathbb{W}^{s,u}_{cj,tT}$
High vs Medium-Low (benchmark)	1.14%	1.34%	-0.20%	-0.19%
(1) Not averaged 1995-2005	1.32%	1.48%	-0.15%	-0.20%
(2) 1996-2006	1.10%	1.32%	-0.23%	-0.19%
(3) LSCV Bandwidths	0.95%	0.95%	0.00%	-0.24%
(4) Cross-country constant technology	1.15%	1.90%	-0.07%	-0.24
Avg	1.14%	1.43%	-0.13%	-0.22%
(5) High-Medium vs Low	2.50%	3.33%	-0.83%	0.43%
(6) Accounting for Endogeneity	2.45%	3.82%	-1.38%	-0.19%

$\Delta\ln\widehat{MRTS}^{s,u}_{cj,tT}$ = Estimated Marginal Rate of Technical Substitution (growth rate); $\Delta\ln\widetilde{\mathbb{B}}^{s,u}_{cj,tT}$ = Counterfactual Skill Biased Technical Change Effect (growth rate); $\Delta\ln\widetilde{\mathbb{F}}^{s,u}_{cj,tT}$ = Counterfactual Factor Accumulation Effect (growth rate): see decomposition in Equation (23). $\Delta\ln\mathbb{W}^{s,u}_{cj,tT}$ = Wage premium growth rate. Unweighted averages reported.

Table 8: Economic (in)efficiency analysis: markers and convergence.

	(1)	(2)	(3)	(4)
Economic Inefficiency 1995	-0.476***	-0.488***	-0.503***	-0.506***
	(0.04)	(0.04)	(0.04)	(0.04)
Minimum Wage Setting * IT Exposure			0.815**	1.444***
			(0.32)	(0.44)
Union Density * IT Exposure				0.129**
				(0.05)
Constant	0.120***	0.087*	0.047	-0.028
	(0.01)	(0.05)	(0.05)	(0.06)
Sector controls	N	Y	Y	Y
N	902	902	582	573
R^2	0.30	0.32	0.49	0.50

Dep. Variable: change in Economic Inefficiency (i.e., $\widehat{\Theta}^{s,u}_{cj,tT}$) computed through Equation (29) (increasing values denoting decreasing efficiency);
IT Exposure is a sector-specific measure of IT employment change;
Minimum Wage Setting and Union Density are country-specific variables;
Bootstrapped standard errors in parentheses;
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: Counterfactual analysis: wage inequality under full adjustment of wages to MP.

	1995 (actual)	2005 (actual)	2005 (counterfactual)
Gini	0.78	0.76	0.86
Atkinson (adversion parameter=1)	0.86	0.81	0.94
Decile range (9th vs 1st decile)	95.45	72.69	207.8
Interquartile range (75th vs 25th %ile)	33.55	25.34	32.90

Figure 1: Change in wage ratio versus change in labor ratio: growth rates, 1995-2005 (country averages).

WIOD Data. High to medium-low skilled ratios. Wages computed by dividing shares in labor compensation by shares in hours worked (i.e., hourly wage ratios).

Figure 2: Estimated MRTS and confidence intervals at 95% (1995).

Estimated Marginal Rate of Technical Substitution: $\Delta \ln \widehat{MRTS}^{s,u}_{cj,tT}$ = see Equation (21).

Figure 3: SBTC and MRTS growth rates, 1995-2005 (country averages). WIOD data.

$\Delta \ln \widehat{MRTS}^{s,u}_{cj,tT}$ = Estimated Marginal Rate of Technical Substitution (growth rate); $\Delta \ln \widetilde{\mathbb{B}}^{s,u}_{cj,tT}$ = Counterfactual Skill Biased Technical Change Effect (growth rate): see decomposition in Equation (23).

Figure 4: Wage ratio versus estimated MRTS, SBTC and FA: growth rates, 1995-2005 (country averages).

$\Delta \ln \widehat{MRTS}^{s,u}_{cj,tT}$ = Estimated Marginal Rate of Technical Substitution (growth rate); $\Delta \ln \widetilde{\mathbb{B}}^{s,u}_{cj,tT}$ = Counterfactual Skill Biased Technical Change Effect (growth rate); $\Delta \ln \widetilde{\mathbb{F}}^{s,u}_{cj,tT}$ = Counterfactual Factor Accumulation Effect (growth rate): see decomposition in Equation (23). $\Delta \ln \mathbb{W}^{s,u}_{cj,tT}$ = Wage premium growth rate. Bisector lines reported.

Figure 5: Wage ratio versus estimated MRTS, SBTC and FA: growth rates, 1995-2005 (sector averages).

$\Delta \ln \widehat{MRTS}^{s,u}_{cj,tT}$ = Estimated Marginal Rate of Technical Substitution (growth rate); $\Delta \ln \widetilde{\mathbb{B}}^{s,u}_{cj,tT}$ = Counterfactual Skill Biased Technical Change Effect (growth rate); $\Delta \ln \widetilde{\mathbb{F}}^{s,u}_{cj,tT}$ = Counterfactual Factor Accumulation Effect (growth rate): see decomposition in Equation (23). $\Delta \ln \mathbb{W}^{s,u}_{cj,tT}$ = Wage premium growth rate. Bisector lines reported.

Figure 6: Country-sector convergence in economic inefficiency: 1995-2005 growth rates versus initial (1995) values.

Economic inefficiency expressed as absolute value of the discrepancy between estimated MRTS and wage ratio ($\left|\ln \widehat{\tau}^{s,u}_{cj,t}\right|$). Growth rate (see Equation (29)): $\widehat{\Theta}^{s,u}_{cj,tT} = \left|\ln \widehat{\tau}^{s,u}_{cj,2005}\right| - \left|\ln \widehat{\tau}^{s,u}_{cj,1995}\right|$. Initial value: $\left|\ln \widehat{\tau}^{s,u}_{cj,1995}\right|$. Univariate OLS regression line reported.

Figure 7: Actual versus counterfactual (full adjustment of wages to MPs) wage distribution: 2005.

In the counterfactual distribution wages are assumed to vary in the same proportion of marginal productivity in each country-sector, in both the skilled and the unskilled labor market (see Section 6.2).

APPENDIX – NOT INTENDED FOR PUBLICATION

A The empirical methodology in details

Nonparametric kernel regression is becoming an increasingly popular method of estimation in applied economic milieus. The main perceived benefit is that it allows for consistent estimation when the underlying functional form of the regression function is unknown. While this is true, there are many other benefits which may prove to be just as useful in our context.

Nonparametric kernel methods (basics). Arguably the most popular regression model in the growth empirics literature is the linear parametric model

$$y_i = \alpha + \beta x_i + \varepsilon_i, \qquad\qquad i = 1, 2, \ldots, n, \tag{31}$$

where y_i is our response (in this case output growth), x_i is a vector of q regressors, α and β are unknown parameters to be estimated and ε_i is the additive (mean zero) random disturbance. Consistent estimation of this model requires that all relevant regressors are included in x_i (and that they are uncorrelated with ε_i) and the functional form is correctly specified. However, when either of these two assumptions do not hold, the estimates the model produces will most likely be inconsistent. While non-linear functional forms are possible in a parametric framework, the data generating process still must be assumed *a priori*.

Nonparametric kernel methods have the ability to alleviate many of the restrictive assumptions made in the parametric framework. Consider the nonparametric regression model

$$y_i = m(x_i) + u_i, \quad i = 1, 2, \ldots, n, \tag{32}$$

where $m(\cdot)$ is an unknown smooth function and the remaining variables are the same as before. Here, $m(\cdot)$ is interpreted as the conditional mean of y given x. Note that in the (linear) parametric setting above, it is implicitly assumed that $E(y_i|x_i) = \alpha + \beta x_i$. Further note that the linear model is a special case of our nonparametric estimator and thus, if the true data generating process is indeed linear, then the nonparametric estimator will give results consistent with that model.

One popular method for estimation of the unknown function is by local-constant least-squares (LCLS) regression. The LCLS estimator of the conditional mean function is given as

$$\widehat{m}(x) = \frac{\sum\limits_{i=1}^{n} y_i \prod\limits_{s=1}^{q} K\left(\frac{x^{si}-x^s}{h^s}\right)}{\sum\limits_{i=1}^{n} \prod\limits_{s=1}^{q} K\left(\frac{x^{si}-x^s}{h^s}\right)}, \tag{33}$$

where $\prod\limits_{s=1}^{q} K\left(\left(x^{si}-x^s\right)/h^s\right)$ is the product kernel and h^s is the smoothing parameter (bandwidth) for a particular regressor x^s (see Pagan and Ullah, 1999). The intuition behind this estimator is that it is simply a weighted average of y_i. It is also known as a local average, given that the weights change depending upon the location of the regressors. We estimate the conditional mean function by locally averaging those values of the left-hand-side variable which are "close" in terms of the values taken on by the regressors. The amount of local information used to construct the average is controlled by the bandwidth.

Local-linear least-squares (LLLS). While LCLS is undoubtedly the most popular, and widely available, nonparametric regression estimator, recently there has been an enthusiastic use of the local-linear least-squares (LLLS) regression estimator as an alternative to LCLS. The LLLS regression estimator possesses several theoretical and empirical advantages. Theoretically, the LLLS estimator has a simple finite sample bias that the LCLS estimator, being unbiased in the setting where the conditional mean is indeed linear. Moreover, the LLLS estimator possesses greater flexibility near the boundaries of the data. Empirically, the LLLS estimator automatically produces estimates of the conditional mean and the associated derivatives. This is beneficial as numerical derivatives can be noisy and behave poorly depending upon the localness of the surrounding data, something that the LCLS estimator can suffer from.

In short, LLLS performs weighted least-squares regressions around a point x with weights determined by a kernel function and bandwidth vector. Again, more weight is given to observations in the neighborhood of x. This is performed over the range of x and then the unknown function is estimated by connecting the point estimates. An added benefit is that if indeed the true functional form is linear, the LLLS estimator nests the OLS estimator when the bandwidth is very large.

Specifically, the covariates vector x_i in (32) is defined as $x_i = [x_i^C, x_i^D]$, where a distinction between continuous (x_i^C) and discrete (x_i^D) variables is made. We can further decompose x_i^D as $[x_i^o, x_i^u]$ where x^o captures variables that are ordered by nature, and x^u captures variables that have no natural ordering. ε_i is a random error term and N is the total number of observations.

Estimation of (31) requires the construction of the product kernel, which is the product of univariate kernel functions (smoothing functions) for each variable. A different type of kernel function is used for each type of data (continuous, discrete ordered and discrete unordered). The product kernel is written succinctly as:

$$G_{i,x} = \prod_{s=1}^{q_C} K\left(x_{is}, x_s, h_s^C\right) \prod_{s=1}^{q_u} g^u\left(x_{is}^u, x_s^u, \lambda_s^u\right) \prod_{s=1}^{q_o} g^o\left(x_{is}^o, x_s^o, \lambda_s^o\right), \tag{34}$$

where q_C is the number of continuous covariates (in our example $q_C = 1$) and $K\left(x_{is}, x_s, h_s^C\right)$ is the kernel function used for continuous variables with bandwidth h_s^C, q_u is the number of unordered discrete regressors (in our example $q_u = 1$) with $g^u\left(x_{is}^u, x_s^u, \lambda_s^u\right)$ is the kernel function for a particular unordered discrete regressor with bandwidth λ_s^u and q_o is the total number of ordered discrete regressors with $g^o\left(x_{its}^o, x_s^o, \lambda_s^o\right)$ the kernel function for a particular ordered discrete regressors with bandwidth λ_s^o.

The product kernel is then used to construct point-specific weights which are then used to calculate a local average estimator. While many different local estimators can be deployed, they all generally have the form

$$\widehat{m}(x) = \sum_{i=1}^{N} y_i A_{ix}. \tag{35}$$

where A_{ix} is a function of the product kernel; different types of local estimators will produce different forms of A_{ix}. The estimator in Equation (35) is nothing more that a weighted average of output for observations that are close, where closeness is dictated exclusively through the bandwidths used in the construction of the estimator (see Li and Racine, 2007 and Henderson and Parmeter, 2015 for more intuition).

For the continuous regressor we choose the Gaussian kernel function

$$K\left(x_{is}^C, x_s^C, h_s^C\right) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_{ix}^C - x_s^C}{h_s^C}\right)^2}; \tag{36}$$

where the bandwidth ranges from zero to infinity.

A variation of the Aitchison and Aitken (1976) kernel function for unordered categorical regressors is given as

$$g^u\left(x_{is}^u, x_s^u, \lambda_s^u\right) = \begin{cases} 1 - \lambda_s^u & \text{if } x_{is}^u = x_s^u \\ \frac{\lambda_s^u}{d-1} & \text{otherwise} \end{cases}; \tag{37}$$

where the bandwidth is constrained to lie in the range $[0, (d-1)/d]$ and d is the number of unique values the unordered variable will take. For example, for the case where the unordered variable is a traditional "dummy variable", the upper bound will be 0.50.

Finally, the Wang and Van Ryzin (1981) kernel function for ordered categorical regressors is given by

$$g^o\left(x_{is}^o, x_s^o, \lambda_s^o\right) = \begin{cases} 1 - \lambda_s^o & \text{if } x_{is}^o = x_s^o \\ \frac{1-\lambda_s^o}{2}(\lambda_s^o)^{|x_{is}^o - x_s^o|} & \text{otherwise} \end{cases}, \tag{38}$$

where the bandwidth ranges from zero to unity.

To estimate the production function in the paper, we use a local-linear approximation which can be viewed as the equivalent of a local Taylor expansion at any point x^c. That is, for the relationship $y = m(x) + u$, we have data for $(y_1, x_1), (y_2, x_2), \ldots, (y_n, x_n)$ and as such, for each point x_i, we can take a linear Taylor approximation for the point x. From the discussion in the paper, x_i is meant as an observation which is indexed over country, sector and time.

To begin, we consider a Taylor expansion about x for observation i as (note the expansion is only for continuous covariates, x^C)

$$\begin{aligned} y_i &= m(x_i) + u_i \\ &\approx m(x) + (x_i - x)\beta(x) + u_i, \end{aligned}$$

where $(x_i - x)$ is a $1 \times q_C$ vector and $\beta(x)$ is the gradient (column) vector of dimension q_C. By ignoring the higher-order terms and treating $m(x)$ and $\beta(x)$ as parameters, we have

$$y_i = a + (x_i - x)b + u_i.$$

Minimizing a quadratic objective function with respect to a and b gives us

$$\widehat{\delta}(x) = \begin{pmatrix} \widehat{m}(x) \\ \widehat{\beta}(x) \end{pmatrix} = \left[\sum_{i=1}^{n} G_{i,x} \begin{pmatrix} 1 \\ x_i - x \end{pmatrix} \begin{pmatrix} 1, & (x_i - x) \end{pmatrix} \right]^{-1} \sum_{i=1}^{n} G_{i,x} \begin{pmatrix} 1 \\ x_i - x \end{pmatrix} y_i$$
$$= \left(\mathbf{X}'G(x)\mathbf{X}\right)^{-1} \mathbf{X}'G(x)y.$$

where $\widehat{\delta} = (\widehat{a}, \widehat{b})$. $\mathbf{X}$ is a $n \times (1 + q_C)$ matrix with first column of all ones and the remaining columns equal to $x_i - x$. Lastly, $G(x)$ is the diagonal matrix with $G_{i,x}$ as its (i,i) element.

Bandwidth selection. It is believed that the choice of the continuous kernel function matters little in the estimation of the conditional mean (see Härdle, 1990) and that selection of the bandwidths is the most salient factor when performing nonparametric estimation. As indicated above, the bandwidths control the amount by which the data are smoothed. For continuous variables, large bandwidths will lead to large amounts of smoothing, resulting in low variance, but high bias. Small bandwidths, on the other hand, will lead to less smoothing, resulting in high variance, but low bias. This trade-off is well known in applied nonparametric econometrics, and the "solution" is most often to resort to automated determination procedures to estimate the bandwidths. Although there exist many selection methods, we utilize the popular least-squares cross-validation (LSCV) criteria. Specifically, LSCV selects bandwidths which minimize

$$CV\left(h, \lambda^o, \lambda^u\right) = \sum_{i=1}^{n} \left[y_i - \widehat{m}_{-i}(x_i)\right]^2, \tag{39}$$

where $\widehat{m}_{-i}(x_i)$ is the leave-one-out estimator of $m(\cdot)$. The idea of the leave-one-out estimator is that the conditional mean of y_i is estimated without using the observation with the most information, x_i. In this way the bandwidths are selected so that the surrounding observations are providing as much information as possible to assist with smoothing. LSCV is well known to produce bandwidths which are quite small relative to the theoretical optimum and as such, will produce estimates which are highly noisy.

An alternative selection mechanism is AIC_c bandwidth selection (Hurvich et al., 1998). The AIC_c criterion is

$$AIC_c(h) = \ln(\widehat{\sigma}^2) + \frac{1 + tr(\mathbf{H})/n}{1 - (tr(\mathbf{H}) + 2)/n}, \tag{40}$$

where

$$\widehat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} \left[y_i - \widehat{m}(x_i)\right]^2 \quad \text{and} \quad \mathbf{H} = \left(\mathbf{X}'G(x)\mathbf{X}\right)^{-1} \mathbf{X}'G(x).$$

Notice that a leave-one-out estimator for $m(\cdot)$ is not used. This is because the AIC_c criterion penalizes overfitting based on the number of effective parameters used, which is captured by the trace of $\mathbf{H}$. As the bandwidths decrease (fit improves) this trace increases and leads to larger penalties. The empirical results in the paper are derived from bandwidths selected using the criterion in (40).

As an aside, we note that an even simpler bandwidth selection procedure, the "ocular" method, is not appropriate once the number of covariates is larger than two. As the number of regressors exceeds two, visual methods to investigate the fit of the model are cumbersome and uninstructive. With a large dimension for the number of regressors, it is suggested that cross-validation techniques be used as opposed to either ocular or rule-of-thumb methods.

Additional references

- Aitchison, J., & Aitken, C. G. (1976). Multivariate binary discrimination by the kernel method. *Biometrika*, 63(3), 413-420.
- Härdle, W. (1990). *Applied nonparametric regression.* Cambridge University Press.
- Hurvich, C. M., Simonoff, J. S., & Tsai, C. L. (1998). Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. *Journal of the Royal Statistical Society: Series B*, 60(2), 271-293.
- Pagan, A., & Ullah, A. (1999). *Nonparametric econometrics.* Cambridge University Press.
- Wang, M. C., & Van Ryzin, J. (1981). A class of smooth estimators for discrete distributions. *Biometrika*, 68(1), 301-309.

B The "Accounting for Endogeneity" robustness check in details

As discussed in the text, potential concerns over endogeneity abound in the estimation of cross-country production functions. As one avenue to remedy this concern, we discuss a potential approach that builds lagged Markovian productivity into the estimation via a control function approach, similar to that adopted in Gandhi, Navarro and Rivers (2020).

Our strategy consists of assuming that time t's inputs are potentially affected by the observed TFP in $t-1$. This is a reasonable assumption in a country-sector context. Assuming that the TFP shock follows the first order Markov process $z_{cj,t} = h_{cj,t}\left(z_{cj,t-1}\right) + \varepsilon_{cj,t}$, the estimating production function can be written as

$$y_{cj,t} = m_{cj,t}(\boldsymbol{x}_{cj,t}) + h_{cj,t}(z_{cj,t-1}) + \varepsilon_{cj,t}, \tag{41}$$

where $\boldsymbol{x}_{cj,t} = (k_{cj,t}, s_{cj,t}, u_{cj,t})$ and $z_{cj,t-1}$ is the lagged productivity shock. Without knowledge of this productivity shock we cannot estimate our production function.

$h_{cj,t}\left(z_{cj,t-1}\right)$ can be retrieved from the estimation of the production function in $t-1$, as the estimated residual. Hence, (41) becomes

$$y_{cj,t} = m_{cj,t}(\boldsymbol{x}_{cj,t}) + h_{cj,t}(y_{cj,t-1} - \hat{m}_{cj,t-1}(\boldsymbol{x}_{cj,t-1})) + \varepsilon_{cj,t}. \tag{42}$$

The difficult is in ensuring that the production function inside of $h_{cj,t}(\cdot)$ is identical to that appearing on its own. To do this we borrow from the insights of the nonparametric instrumental variable estimator of Su and Ullah (2008). This estimator works as follows. For an initial time period, t_0, the production function in Equation (41) is estimated and the residuals $\widehat{\varepsilon}_{cj,t_0}$ are obtained. These residuals are then included as a covariate in the production function in Equation (41) (not explicitly accounting for the presumed additive separability):

$$y_{cj,t} = g_{cj,t}(\boldsymbol{x}_{cj,t}, \widehat{\varepsilon}_{cj,t_0}) + \varepsilon_{cj,t}, \tag{43}$$

This model is estimated using local-linear least-squares as described above and then counterfactual estimation is performed to separate out the two component functions $m_{cj,t}(\cdot)$ and $h_{cj,t}(\cdot)$ using

$$\widehat{m}_{cj,t}(\boldsymbol{x}_{cj,t}) = n^{-1}\sum_{j=1}^{n} \widehat{g}_{cj,t}(\boldsymbol{x}_{cj,t}, \widehat{\varepsilon}_{cj,t_0}).$$

From here, new residuals are calculated in time period t_0, $\widehat{\varepsilon}_{cj,t_0}$ and the process is iterated until convergence. We stop after a tolerance of 0.0001 has been achieved, where our tolerance is based on the estimated function evaluated at all points using squared error.

C Estimated MP in details

We hereby report the estimated MPs, together with plots focusing on time consistency and (non) evidence of decreasing MP.

To capture the dynamics in the two markets, it is useful to compare the growth rate of the MP with that of the wage level and the used amount of labor in the two markets, as in Figure C.1. In the skilled labor market, the overall increase in MP (.041 on average) is associated with shrinking (PPP) real wages ($-.31$ on average) and growing used quantity (.53 on average). Interestingly, while the wages of unskilled workers' shrink at a similar rate, a slight overall decrease in their MP is detected ($-.01$ an average), associated with a slight growth in U (.02).

These trends map into increasing economic inefficiency in the use of skilled labor, relative to unskilled labor, during the decade under consideration.

Figure C.2 focuses on the time consistency of our estimated MPs. Indeed, these reveal quite consistent over time, with both MPS and MPU increasing by more than 25% on average. Unskilled labor features a lower median and a higher (much higher than skilled labor) dispersion. Interestingly, we report an average reduction in MPK (-5.6%), associated with a slight increase in dispersion (around 10%).

Figures C.3 and C.4 show that, differently from capital, increasing (either skilled or unskilled) labor is not associated with lower MP, neither in levels (Figure C.3) nor in growth rates (Figure C.4). Thus, no evidence of decreasing MP emerges, for labor, from our estimates. For instance, we find the high-skill share of hours worked to increase from 23.1% to 28.5%, with the low-medium skill share shrinking from 31.3% to 24.2% (mostly driven by low skills). Under the conventional view of decreasing MPs and Hicks-neutral technical change, this would imply a relative reduction in the MRTS between skilled and unskilled labor. Figures C.3 and C.4 show that this is not the case, according to our analysis.

Finally, Figure C.5 reports the distribution of the estimated MP of (in order, from left to right) unskilled labor (MPU), skilled labor (MPS) and capital (MPK) at the start and finish of the period. Again, estimates look quite consistent over time.

Figure C.1: Estimated MP ($(\widehat{m}^{s}_{cj,t})$, $(\widehat{m}^{u}_{cj,t})$) *versus* wage level and used amount of labor: Skilled (S) and Unskilled (U) labor, country-sector growth rates, 1995-2005.

Figure C.2: Estimated MP $((\widehat{m}^s_{cj,t}), (\widehat{m}^u_{cj,t}), (\widehat{m}^k_{cj,t}))$: time consistency.

Figure C.3: Estimated MP $((\widehat{m}^s_{cj,t}), (\widehat{m}^u_{cj,t}), (\widehat{m}^k_{cj,t}))$ versus (S,U,K) input quantity (2005).

Figure C.4: Change in Estimated MP $((\widehat{m}^s_{cj,t})$, $(\widehat{m}^u_{cj,t})$, $(\widehat{m}^k_{cj,t}))$ versus change in the (S,U,K) input quantity (1995-2005).

Figure C.5: Estimated MP $((\widehat{m}^s_{cj,t})$, $(\widehat{m}^u_{cj,t})$, $(\widehat{m}^k_{cj,t}))$ in 1995 (left) and 2005 (right).