Macroeconomic Dynamics (2023), **27**, pp. 1597–1627
doi:10.1017/S1365100522000372

ARTICLE

Aggregate elasticity of substitution between skills: estimates from a macroeconomic approach†

Michał Jerzmanowski and Robert Tamura*

John E. Walker Department of Economics, Clemson University, Clemson, SC, USA
*Corresponding author. Email: rtamura@clemson.edu

Abstract

We estimate the elasticity of substitution between high-skill and low-skill workers using panel data from 32 countries during 1970–2015. Most existing estimates, which are based only on US microdata, find a value close to 1.6. We bring international data together with a theory-informed macro-approach to provide new evidence on this important macroeconomic parameter. Using the macro-approach, we find that the elasticity of substitution between tertiary-educated workers and those with lower education levels falls between 1.7 and 2.6, which is higher than previous estimates but within a plausible range. In some specifications, estimated elasticity is above the value required for strong skill-bias of technology, suggesting strong skill-bias is possible.

Keywords: Elasticity of substitution; high-skill labor; low-skill labor; skill premium; strong skill-bias; endogenous-directed technology

JEL classifications: E24; E25; O11; J31

1. Introduction

Aggregate elasticity of substitution between workers with different skill levels is an important macroeconomics parameter. For example, it is crucial for quantifying the impact of technological and structural changes on the macroeconomy since it determines how changes in labor composition and technology affect relative wages (Acemoglu, 1998, 2002; Acemoglu and Autor, 2011; Krusell et al. 2000). Importantly, with endogenous-directed technological change, productivity growth can have a so-called *strong skill bias*, which means that an increase in the relative supply of skilled workers can—counter to the standard negative supply effect—raise the wages of those workers (and thus amplifies the skill premium). This happens when the larger supply of skilled workers induces the development of technologies that complement skill, thus offsetting the standard supply effect that pushes wages down. Crucially, this can only occur when the elasticity of substitution between skill types is high enough. Related to this is the question of the quantitative impact of public policies, such as education subsidies, on skill acquisition and evolution of earnings inequality. Here too, the elasticity parameter plays a crucial role (Heckman et al. 1998).

The elasticity parameter is also important in understanding international income differences. Jones (2014) develops a generalized human capital approach and, within that framework, analyzes how much of the cross-country income differences can be explained by human capital, as opposed to the unobserved TFP. The crucial parameter turns out to be the elasticity of substitution between workers with different skill levels. Under the traditional estimates (around 1.6), Jones' generalized

†We would like to thank the participants of the Southern Economics Association's Meetings in New Orleans, Simon Gilchrist, the editors, and two anonymous referees for helpful comments and suggestions.

human capital measure can explain most—or even all—of the income differences. However, with higher values, the explanatory power of human capital falls off sharply. This is a similar finding to Jerzmanowski and Tamura (2019), who compute skill-specific productivity levels for a large sample of countries. They show that barriers to technology adoption explain most of cross-country income differences under the assumption that elasticity of substitution between skill levels is 2.6; however, if the elasticity is closer to the existing consensus of 1.6, human capital accumulation plays a much more important role than barriers to technology. More generally, conclusions from macromodels calibrated to explain the large cross-country dispersion of productivity—the so-called *development accounting* studies—usually depend in important ways on the degree to which workers with different skills can be substituted (Klenow and Rodríguez-Clare, 1997). For example, as Caselli and Ciccone (2019) explain, in order to reconcile the fact that rich countries have much larger relative supplies of skilled workers than poor countries with the fact that skill premia are not significantly lower in rich countries, one must accept that the relative productivity of skilled workers is considerably higher in rich countries. How much higher, however, depends on the value of elasticity of substitution, with lower values implying larger relative productivity gaps between rich and poor economies.[1]

There are however some potential issues with the studies that estimate and use the value of the elasticity of substitution. First, the kinds of aggregate quantitative exercises described above are usually meant to shed light on growth and development in a wide cross-section of countries, many of them at diverse levels of development and with vastly different skill compositions of the labor force. This may be problematic since, as pointed out by Jones (2014), the value of the elasticity of substitution between skills used in these studies is usually based on the microevidence almost exclusively from the USA. Additionally, as recently pointed out by Bowlus et al., (2017), micro estimates using data covering an extended period of time may have a measurement problem, since they assume that workers with a given education level supply the same amount of human capital today as they did more than 60 years ago, long before many of the modern technologies, such as IT, have been used in the workplace and the classroom.

We contribute to this literature by estimating the elasticity of substitution using a macropanel data from a *large group of economies* with most observations coming from only the more *recent time periods*. While the economies in our sample are mostly developed, there is a significant degree of variation among them, both in terms of income levels as well as labor force structure (e.g. we have emerging economies like Poland alongside economies like Germany and the USA in our data). Our explicitly macroeconomic approach means that we need to carefully derive our estimating equation from an appropriate macro model and interpret the coefficient estimates in a manner consistent with the theory. To this end, we draw on the directed technological change literature (Acemoglu, 1998, 2002) to develop and an endogenous-directed technology model with international technology diffusion and capital accumulation. Using this model, we derive the appropriate estimating equation and show that—in a cross-country setting with technology diffusion—the elasticity is not a simple inverse of the wage/labor supply regression coefficient. To the best of our knowledge, this is the first paper to use the international data to estimate the elasticity parameter. We do so rigorously basing it on the theory of directed technical change with cross-country technology diffusions and demonstrate that, in the cross-country setting, interpreting the estimates through the lens of the theory is necessary to recover sensible estimates of elasticity.

We estimate this equation using data from the EU KLEMS Growth and Productivity Accounts panel data set (EU KLEMS project on Growth and Productivity in the European Union, 2018), a detailed database of industry-level measures of output, inputs, and productivity for 30 European countries (most of Europe plus Japan, South Korea, Australia, and the USA) for the period from 1970 to 2015 (with 90% of observations after 1980). We find that the elasticity of substitution between tertiary-educated workers and those with lower education levels likely falls within the range of 1.7 and 2.6, which is higher than previous estimates but within a plausible range. In most of our regressions, the estimated elasticity falls short of the value required for strong skill-bias of technology; however, in some specifications it is above that level (or at least the 95% confidence interval includes strong bias), suggesting strong skill-bias is possible. When we restrict our attention to a narrower definition of low-skilled workers (low-skill group only instead of low- and middle-skilled groups together), the findings point even more strongly in the direction of higher elasticity and a possibility of a strong skill-bias. Finally, we also estimate our model using disaggregated data for nine major sectors. They are similar to the aggregate results and well within the plausible range but in many cases have slightly higher values. Interestingly, in some instances, most notably in the finance and the education sectors, there is a quite pronounced indication of a strong skill bias.

Our conclusions for this exercise are as follows. First, using a macropanel data of a fairly diverse group of countries over an extended time period, the elasticity estimates are largely in line with existing estimates obtained using very different data sets. We feel this is encouraging and strengthens the confidence we have in the range of plausible elasticity values. Second, we demonstrate that when using cross-country data, accounting for technology diffusion and endogeneity of the direction of technological change (what we dub the "macro-approach") are important for correctly interpreting the elasticity estimates. We feel that further work using international data to investigate elasticity of substitution between skills—for example, by relaxing our assumption of balanced growth path or use of better instrumental variables—should continue to pay close attention to the underlying theoretical structure. Finally, the fact that some of our results point in the direction of higher elasticities is also consistent with the most recent findings using more disaggregated data, further strengthening their plausibility.

The paper is organized as follows. Section 1 discusses the canonical approach and the resulting estimating equation used in the literature based on the US microdata. It then describes our macro-approach and derives the estimating equation and its interpretation appropriate for the cross-country context. Section 2 discusses our data and estimation, while section 3 presents the baseline results, along with some robustness checks. Section 4 concludes.

2. Theoretical approach

In this section, we describe the standard theory behind most existing elasticity estimates. We then show how a model with endogenous and directed technological progress and cross-country diffusion of knowledge leads to a different interpretation of the otherwise conventional "elasticity" regression. The detailed derivations of the model are left to the appendix; in this section, we focus only on the key equations.

2.1 The canonical approach

The traditional approach to estimating the elasticity of substitution between workers with different skill sets is based on the constant elasticity of substitution production (CES) function. Most studies start with a CES production function with two distinct categories of labor: high-skilled and low-skilled.

$$Y = \{(A_H H)^{\frac{\sigma-1}{\sigma}} + (A_L L)^{\frac{\sigma-1}{\sigma}}\}^{\frac{\sigma}{\sigma-1}} \tag{1}$$

where H (L) denotes the quantity of skilled (unskilled) labor, A_H and A_L are skill-specific productivity levels, and σ is the elasticity of substitution.

Under competitive behavior by firms, profit maximization leads to first-order conditions equating the marginal product of labor to wages for each type of labor. Taking the ratio of these conditions leads directly to an expression relating relative wages to relative supplies

$$\log\left(\frac{w_H}{w_L}\right) = \frac{\sigma-1}{\sigma}\log\left(\frac{A_H}{A_L}\right) - \frac{1}{\sigma}\log\left(\frac{H}{L}\right). \tag{2}$$

This equation can be transformed into a regression by appending an error term, and, given data on relative wages and relative skill supplies, it can be estimated. The value of elasticity σ is

https://doi.org/10.1017/S1365100522000372 Published online by Cambridge University Press

then backed out as the inverse of the estimate of the coefficient on log relative supplies. Because technological progress (and perhaps institutional change) is likely to imply changing growing productivity levels, the term A_H/A_L is usually proxied by a linear or quadratic trend.

In practice, this equation is usually estimated using microdata: wages and supplies are constructed from CPS or Census data and aggregated up to the country or state levels. This is the approach in the seminal paper by Katz and Murphy (1992), who use CPS data over the period 1963–87. Ciccone and Peri (2005) also use CPS data (between 1950 and 1990) but construct their wage and supply measures at the state level in order to exploit an instrumental variable approach. These studies usually find values of the elasticity parameter between 1.4 and 1.6, in line with earlier literature on the topic (Johnson, 1970). However, when quadratic or even more flexible terms are used to proxy for the technology term, it is not unusual to find estimates of σ in excess of two. Similarly, more recent data seem to favor higher elasticity. Acemoglu and Autor (2011) find that extending the Katz and Murphy sample to 2008 yields an elasticity estimate as high as 2.9.[2] Nevertheless, the consensus seems to remain that the value is somewhere in the vicinity of 1.6.

A recent paper by Bowlus et al., (2017) argues that there is a potential problem with the measurement of skill supplies in the literature estimating the canonical model. Specifically, aggregating workers using average wages as weights implicitly assumes the same quantity of human capital is supplied by workers of a given education level, regardless of whether they obtained their education in 1960 or 2010. When they incorporate a correction, based on calculating quantities using aggregate wage bill and price (wage) information, they report considerably higher estimates of elasticity of substitution (around 3.5). Our data, detailed below, cover the period 1970–2015 but most of the observations (90%) come from the period after 1980, making the problem of changes in the quality of education over time potentially less pertinent.

2.2 The macro-approach

Because we wish to estimate the elasticity of substitution using aggregate macrodata from a panel of countries, our starting point is a macroeconomic model where the rate and direction of technological progress is endogenous and technologies are allowed to diffuse across economies, as they surely do in practice. The model, based on the seminal work on directed technological progress by Acemoglu (1998, 2002), was developed in Jerzmanowski and Tamura (2019); here we give a brief sketch of its main parts, with detail relegated to the appendix.

The final output in our economy is produced by competitive firms combining two intermediate good inputs according to a CES aggregator. The two varieties of intermediate inputs come from two distinct intermediate sectors and differ in terms of the labor input required to produce them. An intermediate sector combines physical capital and labor of either high-skill or low-skill according to a Cobb-Douglas production function. Denoting the intermediate good output by Y_i (where $i = H, L$ stands for either a high-skill or a low-skill sectors), the (reduced form) final output is given by

$$Y = \left\{ \left(K_H^{1-\beta} (A_H H)^{\beta} \right)^{\frac{\varepsilon-1}{\varepsilon}} + \left(K_L^{1-\beta} (A_L L)^{\beta} \right)^{\frac{\varepsilon-1}{\varepsilon}} \right\}^{\frac{\varepsilon}{\varepsilon-1}}, \tag{3}$$

where H (L) is the endowments of high(low)-skill labor, A_H (A_L) is the endogenous productivity of high(low)-skill labor, and K_H (K_L) is the amount of physical capital used by high(low)-skill workers.

In equilibrium, the relative wage of the two types of workers is given, just as in the canonical model, by

$$\frac{w_H}{w_L} = \left(\frac{A_H}{A_L} \right)^{\frac{\sigma-1}{\sigma}} \left(\frac{H}{L} \right)^{-\frac{1}{\sigma}}, \tag{4}$$

where $\sigma = 1 + (\varepsilon - 1)\beta$ is the elasticity of substitution between worker types.

https://doi.org/10.1017/S1365100522000372 Published online by Cambridge University Press

Notice in equation (4) that an increase in H/L has a direct effect of reducing the relative skilled wage through the standard supply effect, which is used in the canonical approach to identify the elasticity parameter. However, following the work of Acemoglu (1998, 2002), we make the skill-specific productivity levels (A_H and A_L) endogenous by assuming that profit-driven innovators supply new technologies to each sector. The resources devoted to innovation for each sector, and, as a result, the rate of growth of productivity depends on the sector's size (as measured by the supply of workers in each skill category). This means that the term A_H/A_L in equation (4) depends on the relative supply H/L. In addition, our model includes technology diffusion, whereby innovators in every country benefit from the world stock of knowledge. In the appendix, we show that along the balanced growth path, the relative level of productivities is given by

$$\frac{A_H}{A_L} = \left(\frac{\eta_H}{\eta_L}\right)^{\frac{\sigma}{1+\varphi\sigma}} \left(\frac{H}{L}\right)^{\frac{\sigma-1}{1+\varphi\sigma}} \left(\frac{A_H^W}{A_L^W}\right)^{\frac{\varphi\sigma}{1+\varphi\sigma}}, \tag{5}$$

where φ measures the strength of technology diffusion, η_i is the efficiency of the innovation process aimed at sector $i = H, L$, and A_i^W denotes the world technology frontier for sector $i = H, L$.[3] Substituting this expression into equation (4) yields

$$\frac{w_H}{w_L} = \left(\frac{\eta_H}{\eta_L}\right)^{\frac{\sigma-1}{1+\sigma\varphi}} \left(\frac{H}{L}\right)^{\frac{\sigma-2-\varphi}{1+\sigma\varphi}} \left(\frac{A_H^W}{A_L^W}\right)^{\frac{\varphi(\sigma-1)}{1+\sigma\varphi}}, \tag{6}$$

Notice the effect of an increase in relative supply of skilled workers (H/L) has two effects on their relative wages. In addition to the direct supply effect of H/L on wages (equation (4)), the supply increase also raises the relative productivity A_H/A_L if the term $\frac{\sigma-1}{1+\varphi\sigma}$ in equation (5) is positive. When this increase in relative productivity is strong enough to offset the supply effect and leads to an increase in the relative equilibrium wage of skilled workers, we—following Acemoglu's terminology—refer to it as (relative) *strong skill-bias*. Clearly, the *strong skill-bias* is present in equilibrium as long as

$$\sigma > 2 + \varphi \tag{7}$$

which reduces to $\sigma > 2$, a result familiar from Acemoglu (2009), when there is no technology diffusion ($\varphi = 0$). This means that with a sufficiently higher substitutability between skills, an increase in (relative) supply of skilled workers (an increase in *market size* for skill-biased technology) induces an increase in (relative) productivity of these workers (A_H/A_L) that is large enough to offset the usual negative effect on their marginal product (the term $(H/L)^{-\frac{1}{\sigma}}$ in equation (4)). As a result, the (relative) wages of skilled workers rise. Notice that the presence of international technology diffusion ($\varphi > 0$) implies a higher value of σ is required for a strong bias to exist. This happens because the presence of technology diffusion means that some of the relative productivity (A_H/A_L) changes come from the world technology frontier and are independent of domestic market size (equation (5)). For the effect coming from just the domestic market to be large enough, the elasticity of substitution must be even higher than in the absence of diffusion.

3. Data and estimation

Taking logs of equation (6) leads to a linear relationship

$$\log\left(\frac{w_H}{w_L}\right) = \frac{\sigma-1}{1+\sigma\varphi}\log\left(\frac{\eta_H}{\eta_L}\right) + \frac{\sigma-2-\varphi}{1+\sigma\varphi}\log\left(\frac{H}{L}\right) + \frac{\varphi(\sigma-1)}{1+\sigma\varphi}\log\left(\frac{A_H^W}{A_L^W}\right),$$

which—after appending an error term and approximating the skill-bias of world technology frontier A_H^W/A_L^W and the evolution of relative innovation efficiencies η_H/η_L with functions of

https://doi.org/10.1017/S1365100522000372 Published online by Cambridge University Press

time t (in practice, we use linear, quadratic, and country-specific trends)—becomes a regression equation[4]

$$\log\left(\frac{w_H}{w_L}\right)_{it} = \gamma_0 + \gamma_1\,\mu(t) + \gamma_2 \log\left(\frac{H}{L}\right)_{it} + \varepsilon_{it}, \tag{8}$$

to be estimated using data on relative wages and relative labor supplies, where γ_0 and $\gamma_1\mu(t)$ stand for the parameters—later modeled as time trend polynomials or year fixed effects—that capture the movement of the time-varying but unobservable factors in the equation: the relative world technology frontier ($\frac{A_H^W}{A_L^W}$) and the relative innovation efficiency parameters ($\frac{\eta_H}{\eta_L}$). The technology frontier is common to all countries, whereas the innovation efficiency may depend on country-specific factors such as laws, institutions, and regulations. We therefore explore variety of specifications, including common and country-specific trends, when controlling for these factors. Our key object of interest, the value of the elasticity of substitution between high- and low-skilled workers, can then be recovered from the estimate of γ_2 using

$$\gamma_2 = \frac{\sigma - 2 - \varphi}{1 + \sigma\varphi}, \tag{9}$$

and its standard errors can be computed using the Delta method (Greene, 2003).

Notice that this is the same regression as the one estimated in most papers on the elasticity of substitution between labor of different skill levels. However, the structural interpretation of the coefficient on the relative skill supplies (γ_2) is different.[5] In the presence of directed technological change and technology diffusion, this coefficient is not the inverse of the elasticity of substitution σ, as was the case in the canonical approach, and additionally, it depends on the diffusion parameter φ. This means that in order to recover the value of the elasticity parameter from the estimate of the coefficient on log relative supply, we need to use equation (9) and we need to know the value of φ. In Jerzmanowski and Tamura (2019), we work with a calibrated version of the above model in order to compute skill-specific productivity levels for a large sample of countries. We show that given plausible values of other model parameters, a value 0.5 for φ produces plausible dynamic behavior of the model (specifically the rate of convergence to the balanced growth path matches estimates found in the literature). We, therefore, choose to use $\varphi = 0.5$ as our preferred value. However, since this parameter does not have a generally agreed-upon value, we also calculate the value of elasticity under other plausible magnitudes of the diffusion rate.

We estimate σ using equation (8) and data from the EU KLEMS Growth and Productivity Accounts panel data set (EU KLEMS project on Growth and Productivity in the European Union, 2018; Jäger, 2018; O'Mahony and Timmer, 2009). This is a detailed database of industry-level measures of output, inputs, and productivity for 28 European countries, Japan, South Korea, Australia and the USA for the period from 1970 to 2015. The data, which come at various levels of sectoral disaggregation depending on the time period, provide information on the share of hours worked and wages, broken down into three skill groups: low skill (less than high school degree), medium skill (high school degree) and high skills (college). These come from various survey sources, including the European Labour Force Survey for many E.U. countries, and the Census and CPS for the USA Because of the differences in the definition of medium- and low-skilled workers across countries and over time, we decided to combine these two categories together to form a *lower skill* group. The definition of high-skill workers is fairly uniform over time and across countries and almost always includes individuals with some college and above (Timmer et al. 2007).[6] For our main results, we use the country aggregates, designated as *total economy* in EU KLEMS, which sums all sectors in each country. We transform the values of the share of hours worked and the share of the wage bill into relative hours and wages for the purpose of our empirical specification. For our combined middle- and low-skill group, we follow Katz and Murphy (1992) and compute the average wages of the group in each country as wages weighted by average hours supplied over

Table 1. Summary statistics

Variable	Mean	Std. Dev.	Min	Max
Share of high-skill hours	0.24	0.09	0.06	0.49
Share of middle-skill hours	0.48	0.14	0.01	0.80
Share of low-skill hours	0.29	0.16	0.04	0.78
Share of high-skill wages	0.33	0.11	0.10	0.63
Share of middle-skill wages	0.45	0.13	0.05	0.73
Share of low-skill wages	0.22	0.15	0.02	0.68
High-skill wage/low-skill wage	1.71	0.31	0.88	2.81
High-skill wage/middle-skill wage	1.56	0.31	0.33	2.91
High-skill wage/low-skill wage	2.14	0.81	1.14	9.16
$N = 642$				

the entire sample of the skill group. We then weigh the hours for each country-year cell by those average wages to create a supply of lower-skilled workers. As we already mentioned, Bowlus et al., (2017) argue that there is a potential problem with the measurement of changes in skill supplies when using this approach since it implicitly assumes the quantity of human capital supplied by workers of a given education level is identical, regardless of whether they obtained their education in the distant past or more recently. Our sample covers the period 1970–2015 but 90% of the observations come from the period after 1980, making this problem potentially less pertinent.

The EU KLEMS labor data come in several versions. The original release, published in 2007, contained data from 1970 to 2005 (with shorter series for some countries). Subsequent releases contained data for the original countries and some new ones for the period 2006–2017. In our analysis, we used combined data for the years 1970–2015.[7] The potential for changes in definitions of skill groups over time leads the database authors to recommend the following strategy for ensuring comparability of the hour and wage shares over time: compute a time series of *annual growth rates* and combine them with the *most recent level* information (2015 for us) to compute prior years' levels. We follow this recommendation. It turns out that using the raw data produces very similar results, which are available upon request. To further ensure any changes in skill definitions are not affecting our analysis, we perform our estimation using only the original pre-2005 sample.

Finally, we note that equation (8) holds along the balanced growth path for a constant supply of labor in the different skill categories and constant growth rate of skill-specific productivity levels. Of course, this is not exactly true in our data set but if most of our countries are sufficiently close to their respective balanced paths and the changes in relative skill supply shifts they experience are not too large and rapid, the equation will provide a sufficiently good approximation to the true evolution of relative wages. Estimating a fully dynamic model to account for adjustments off the balanced growth path is left for future research.

Table 1 contains the summary statistics of the hours and wage *shares* for each of the three skill categories, as well as the calculated *relative wages* and *relative labor supplies* we calculate based on the shares. Our main results are obtained using regressions of relative wages on relative supplies of high-skilled to lower-skilled (low and middle skill combined). We present robustness analysis using relative values of high-skilled to middle-skilled only.

4. Results

This section presents and discusses the results of estimating equation (8) using the EU KLEMS data on wages and labor supplies of high-skilled and lower-skilled workers groups. We use ordinary least squares, fixed effects, instrumental variables (where we instrument H/L with its lagged

Table 2. College vs. lower-skilled; OLS

	1	2	3	4	5	6
log (H/L)	−0.001	0.001	−0.019	−0.013	−0.036	0.003
	(0.050)	(0.049)	(0.063)	(0.059)	(0.063)	(0.049)
Country effects	No	No	No	No	No	No
Trend	Yes	Yes	Yes	Yes	Yes	No
Trend squared	No	Yes	No	Yes	Yes	No
Country trend	No	No	Yes	Yes	Yes	No
Country trend Sq	No	No	No	No	Yes	No
Year Eeffects	No	No	No	No	No	Yes
σ	2.50	2.50	2.46	2.47	2.42	2.51
s.e.	(0.113)	(0.111)	(0.140)	(0.131)	(0.137)	(0.110)
95% conf. int.	[2.27, 2.72]	[2.28, 2.72]	[2.18, 2.74]	[2.21, 2.73]	[2.15, 2.70]	[2.29, 2.73]
p($\sigma < 2.5$)	0.51	0.49	0.62	0.59	0.71	0.47
σ'	839.00	−888.02	53.35	77.11	28.08	−318.52
R^2	0.00	0.00	0.44	0.47	0.72	0.00
n	642	642	642	642	642	642

NOTES: Ordinary least squares estimates of $\log(w_H/w_L)_{it} = \gamma_0 + \gamma_1\,\mu(t) + \gamma_2 \log(H/L)_{it} + \varepsilon_{it}$, using high-skilled workers as H and low- and middle-skilled workers for L. Standard errors clustered at the country level. "Macro-"elasticity of substitution σ calculated based on $\gamma_2 = (\sigma - 2 - \varphi)/(1 + \sigma\varphi)$ with $\varphi = 0.5$. Standard error of σ calculated using the Delta method. $p(\sigma < 2.5)$ denotes the probability that the true "macro-"elasticity is less than the strong skill-bias threshold. Conventional elasticity σ' calculated as $1/\gamma_2$.

values), and system GMM estimators, and in each case, we explore different ways of proxying for the changes in the world technology frontier skill bias (linear, quadratic, and country-specific trends).

For each estimation method and time trend specification, we report the point estimates and the associated standard errors of γ_2 from equation (8); the implied estimate of the elasticity of substitution between high- and low-skill labor σ and its associated two standard deviation interval, using standard errors calculated using the Delta method; and the elasticity of substitution that would be implied under the canonical interpretation of γ_2 (denoted by σ'), which ignores directed technology change and cross-country technology diffusion and computes the elasticity as an inverse of γ_2 estimate.

We start with our preferred estimates of the elasticity of substitution based on the value of the diffusion parameter $\varphi = 0.5$. We then show how the estimates would be affected by imposing a different rate of technology dissemination, choosing a different lower skill definition, or restricting our sample.

4.1 Baseline estimates

Table 2 below shows the results of OLS estimation. When only a linear or quadratic time trend is included (columns 1 and 2), the point estimate of the elasticity of substitution is 2.5, which is higher than most results in the literature but is in the plausible range and, in fact, not far from some of the recent findings (Acemoglu and Autor, 2011) or earlier estimates that used more flexible ways to control for the time trend (Katz and Murphy, 1992). The point estimate is similar if we use a more flexible year effects instead of the trend polynomials (last column). Importantly, note that if were to follow the canonical model's interpretation of the coefficient γ_2 and compute the elasticity as its inverse, we would get obtain a highly implausibly values. The reason is simply that according to the data, the increase in relative supply of skilled workers has zero effect on their relative wage. In our macrosetting, this does not require a vastly implausible elasticity but instead a high but

https://doi.org/10.1017/S1365100522000372 Published online by Cambridge University Press

Table 3. High- vs. lower-skilled; Country fixed effects

	1	2	3	4	5	6
log (H/L)	-0.355^{***}	-0.356^{***}	-0.313^{***}	-0.299^{***}	-0.151	-0.336^{***}
	(0.106)	(0.106)	(0.082)	(0.107)	(0.114)	(0.114)
Country effects	Yes	Yes	Yes	Yes	Yes	Yes
Trend	Yes	Yes	Yes	Yes	Yes	No
Trend squared	No	Yes	No	Yes	Yes	No
Country trend	No	No	Yes	Yes	Yes	No
Country trend Sq	No	No	No	No	Yes	No
Year effects	No	No	No	No	No	Yes
σ	1.82	1.82	1.89	1.92	2.18	1.85
s.e.	(0.172)	(0.172)	(0.137)	(0.183)	(0.221)	(0.188)
95% conf. int.	[1.48, 2.17]	[1.48, 2.17]	[1.62, 2.17]	[1.55, 2.28]	[1.74, 2.63]	[1.48, 2.23]
p($\sigma < 2.5$)	1.00	1.00	1.00	1.00	0.90	1.00
σ'	2.81	2.81	3.19	3.35	6.62	2.98
R^2	0.14	0.16	0.65	0.67	0.76	0.18
n	642	642	642	642	642	642

NOTES: Estimates of $\log (w_H/w_L)_{it} = \lambda_i + \gamma_0 + \gamma_1 \mu(t) + \gamma_2 \log (H/L)_{it} + \varepsilon_{it}$, using high-skilled workers as H and low- and middle-skilled workers for L. Standard errors clustered at the country level. "Macro-"elasticity of substitution σ calculated based on $\gamma_2 = (\sigma - 2 - \varphi)/(1 + \sigma\varphi)$ with $\varphi = 0.5$. Standard error of σ calculated using the Delta method. $p(\sigma < 2.5)$ denotes the probability that the true "macro-"elasticity is less than the strong skill-bias threshold. Conventional elasticity σ' calculated as $1/\gamma_2$.

reasonable value of 2.5. Note that, of course, the value of 2.5 is exactly borderline for the presence of *strong skill bias* since (7) is satisfied as an equality. That is the increase in technology (and thus demand for skills) induced by changing quantity of skilled workers is just powerful enough to offset the usual negative supply effect.

When we allow the time trend to be country-specific, the point estimates of σ decrease slightly to between 2.42 and 2.47 (columns 3–5). These are still higher than the conventionally accepted value of 1.6, but these estimates fall slightly short of the cutoff needed for strong skill bias. The 95% confidence intervals however do include values for which strong bias would be present. The estimates that would be obtained under the standard interpretation, which ignores directed technology and diffusion of ideas, are closer to the realm of plausibility but are much larger than the literature's 1.6 or even our own estimates and, additionally, are highly sensitive to the time trend specification.

Table 3 reports the results of estimating equation (8) with country fixed effects. Here again, the estimates are somewhat lower, in the range of 1.82–2.18 but still considerably higher than 1.6, with values well above 2 comfortably within the confidence interval. And, when the country-specific trend squared is included, the regression implies no significant effect of changes in relative supply on relative wages, that is the skill bias completely offsets the negative supply effect. In this case, the point estimate of elasticity of substitution is firmly above two (and the confidence interval includes value consistent with strong skill bias). The conventionally computed elasticities are much higher, and in three cases, implausibly so.

There is potential for endogeneity bias if a shock to wages induce a response in hours of work supplied on the either intensive or extensive margin (given that our data are at annual frequency, we are less worried about education attainment's response to wages). Unfortunately, we do not have a good candidate for an instrument, but in Table 4 we report the estimate of our model using lagged values of relative labor supplies as instruments for the current level. These results are very similar to those from the fixed effects model, and as before, their conventional interpretation, that is those that ignore technology diffusion cross countries, would lead to estimates of elasticity

https://doi.org/10.1017/S1365100522000372 Published online by Cambridge University Press

Table 4. High- vs. lower-skilled; IV

	1	2	3	4	5	6
log (H/L)	-0.411^{***}	-0.429^{***}	-0.271^{***}	-0.263^{*}	-0.108	-0.400^{***}
	(0.121)	(0.119)	(0.089)	(0.143)	(0.085)	(0.124)
Country effects	Yes	Yes	Yes	Yes	Yes	Yes
Trend	Yes	Yes	Yes	Yes	Yes	No
Trend squared	No	Yes	No	Yes	Yes	No
Country trend	No	No	Yes	Yes	Yes	No
Country trend Sq	No	No	No	No	Yes	No
Year effects	No	No	No	No	No	Yes
σ	1.73	1.71	1.96	1.98	2.27	1.75
s.e.	(0.187)	(0.181)	(0.156)	(0.252)	(0.171)	(0.194)
95% conf. int.	[1.36, 2.11]	[1.34, 2.07]	[1.65, 2.27]	[1.47, 2.48]	[1.93, 2.61]	[1.36, 2.14]
p(σ < 2.5)	0.91	0.91	0.90	0.84	0.79	0.90
σ'	2.43	2.33	3.69	3.80	9.23	2.50
N	537	537	537	537	537	537

NOTES: IV estimates of $\log (w_H/w_L)_{it} = \lambda_i + \gamma_0 + \gamma_1\,\mu(t) + \gamma_2 \log (H/L)_{it} + \varepsilon_{it}$, using high-skilled workers as H and low- and middle-skilled workers for L on the restricted sample (1970–2005). Lagged values of relative labor supplies as instruments for the current level. Standard errors clustered at the country level. "Macro-"elasticity of substitution σ calculated based on $\gamma_2 = (\sigma - 2 - \varphi)/(1 + \sigma\varphi)$ with $\varphi = 0.5$. Standard error of σ calculated using the Delta method. $p(\sigma < 2.5)$ denotes the probability that the true "macro-"elasticity is less than the strong skill-bias threshold. Conventional elasticity σ' calculated as $1/\gamma_2$.

that are—in most specifications—either considerably high (above 3) or even very implausibly large (9.23) and additionally very sensitive to the specification of the time effects. As our final empirical method, we estimate our equation using system GMM estimator (Arellano and Bover, 1995), which, in addition to first-differed equation instrumented with lags of relative supply, uses an equation in levels.[8] Table 5 contains the results of our GMM estimation. The GMM estimates of the elasticity parameter are considerably higher than those using other estimators. The point estimates continue to fall right around the boundary of strong bias, and the 95% confidence intervals always include a region where bias would indeed be of the strong kind. Some of our modeling choices and assumptions make the GMM a potentially appealing approach. For example, the inclusion of time trend polynomials or, even more flexibly, year effects, makes it less likely that idiosyncratic disturbances are correlated across countries (say, due to a common shock to the technology frontier). However, other assumptions may well not hold. The lack of serial correlation in the idiosyncratic part of the error term (net of the country fixed effect) is less likely to be a problem if countries are close to their balanced growth paths, as our approach assumes but does not guarantee this is the case. Thus, the results obtained with the use of internal instruments must be taken with caution. That said, we note that the GMM estimates provide a strong suggestion that the technological change over the last several decades may in a fact have been strongly skill-biased.

We conclude that our estimates, together with our macrointerpretation, which accounts for directed technology change and cross-country idea diffusion, produce a set of estimates of the elasticity of substitution that are higher than those found using US microdata but fall within a plausible range. Importantly, some of the specifications produce point estimates (or at least confidence intervals) consistent with the strong skill-bias of technology. We also note that the elasticity estimates under our macrointerpretation are not overly sensitive to the estimation method and the trend specification. The same cannot be said of the values that would be obtained if we followed the conventional interpretation of the coefficient on relative labor supplies: here, the estimates are highly sensitive to specification and, most of the time, fall outside of a plausible range.

Table 5. High- vs. lower-skilled; system GMM

	1	2	3	4	5	6
log (H/L)	−0.025	−0.022	0.009	0.010	−0.038	−0.020
	(0.064)	(0.062)	(0.057)	(0.056)	(0.062)	(0.061)
Country effects	Yes	Yes	Yes	Yes	Yes	Yes
Trend	Yes	Yes	Yes	Yes	Yes	No
Trend squared	No	Yes	No	Yes	Yes	No
Country trend	No	No	Yes	Yes	Yes	No
Country trend Sq	No	No	No	No	Yes	No
Year effects	No	No	No	No	No	Yes
σ	2.45	2.45	2.52	2.52	2.42	2.45
s.e.	(0.141)	(0.137)	(0.129)	(0.126)	(0.134)	(0.135)
95% conf. int.	[2.16, 2.73]	[2.18, 2.72]	[2.26, 2.78]	[2.27, 2.78]	[2.15, 2.68]	[2.18, 2.73]
p($\sigma < 2.5$)	0.65	0.64	0.44	0.43	0.73	0.63
σ'	40.57	44.49	−107.48	−99.88	26.10	49.48
N	642	642	642	642	642	642

NOTES: System GMM estimates of $\log (w_H/w_L)_{it} = \lambda_i + \gamma_0 + \gamma_1\, \mu(t) + \gamma_2 \log (H/L)_{it} + \varepsilon_{it}$, using high-skilled workers as H and low- and middle-skilled workers for L. Lagged levels of relative labor supplies as instruments for the first-difference equation and lagged differences of relative supplies are used as instruments in the level equation. Standard errors clustered at the country level. "Macro-"elasticity of substitution σ calculated based on $\gamma_2 = (\sigma - 2 - \varphi)/(1 + \sigma\varphi)$ with $\varphi = 0.5$. Standard error of σ calculated using the Delta method. $p(\sigma < 2.5)$ denotes the probability that the true "macro-"elasticity is less than the strong skill-bias threshold. Conventional elasticity σ' calculated as $1/\gamma_2$.

4.2 Robustness

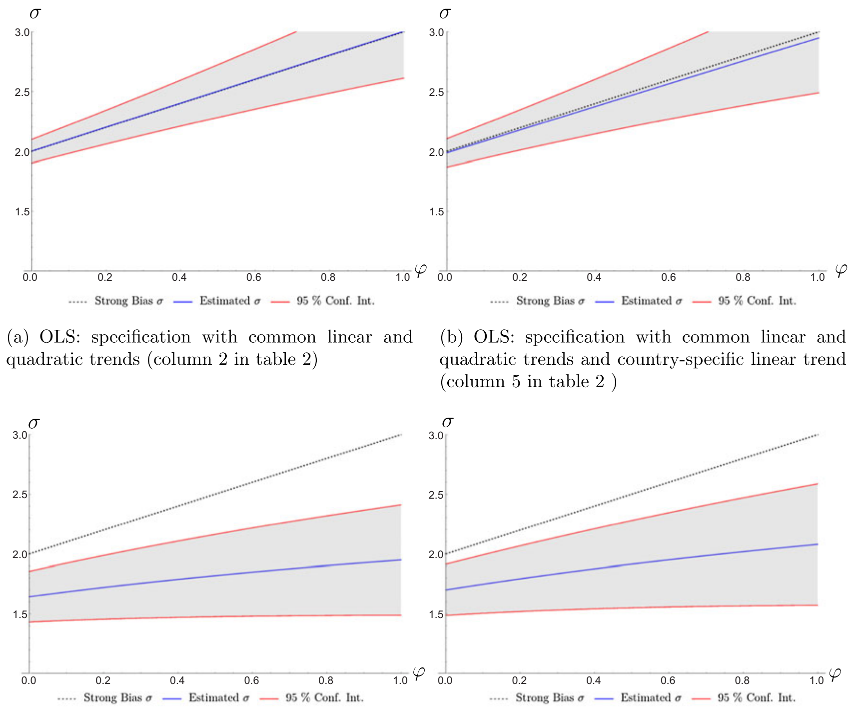
4.2.1. Different rates of diffusion (φ)

In the above tables, we calculated the value of the elasticity of substitution by inverting equation (9) and using our preferred value of the diffusion rate $\varphi = 0.5$. Here we present the estimates of σ (as well as the two standard deviation interval) computed using alternative values of φ. Specifically, in Figure 1, we plot—for OLS and fixed effects specifications—the values of σ against the diffusion rate (in each case using two out of the five specifications reported in the tables). The shaded region corresponds to the two standard deviation band, and the black line is the value of elasticity above which the strong skill bias is present. We repeat this for the IV and GMM estimates in Figure 2.

As can be seen from the graphs, given the estimated regression coefficients, the implied elasticity is increasing in φ, but, of course, so is the cutoff for strong skill bias. In the case of OLS, point estimates are always only slightly below the threshold when a country-specific linear trend is not included (Figure 1(a)) but fall farther from the threshold once this variable is included in the regression (Figure 1(b)). For both OLS specifications, the elasticity consistent with strong bias falls within the two standard deviation interval. Under country fixed effects (Figure 1(c)–1(d)), the point estimates of elasticity are lower and, as a result, the strong skill-bias cutoff falls outside of the two standard deviation regardless of the diffusion rate. The GMM results (Figures 2(c)–2(d)) are similar to the OLS ones and the case of IV (Figures 2(a)–2(b)) falls in between. Without the country-specific trend, strong skill bias is very implausible, regardless of the diffusion value (Figure 2(a)). When a country-specific trend is included, strong skill bias is within the two standard deviation band but only for diffusion rates lower than our preferred value of 0.5 (Figure 2(b)).

We also note that, relative to our preferred results with the diffusion rate of 0.5 presented above, the absence of any technological diffusion ($\varphi = 0$) implies elasticity values closer to those obtained in the literature, while faster diffusion suggests the evidence is consistent with a much greater degree of substitutability between skilled and unskilled workers.

https://doi.org/10.1017/S1365100522000372 Published online by Cambridge University Press



(a) OLS: specification with common linear and quadratic trends (column 2 in table 2)

(b) OLS: specification with common linear and quadratic trends and country-specific linear trend (column 5 in table 2)

(c) FE: specification with common linear and quadratic trends (column 2 in table 3)

(d) FE: specification with common linear and quadratic trends and country-specific linear trend (column 5 in table 3)

Figure 1. Estimates of σ (blue line) and the two standard deviation intervals (shaded) for different values of the diffusion parameter φ based on different specifications under either OLS or fixed effects (FE). The black line denotes the value required for strong skill bias.

4.2.2. Different definition of lower-skilled

In the regressions above, we have combined the low and medium skill groups from the EU KELMS data. We have done this, because the distinction between these two categories is not uniform across countries in the sample (O'Mahony and Timmer, 2009). However, much of the literature estimating the elasticity of substitution between skill types uses two groups: college-educated and high-school-educated workers. Thus, we repeat our analysis, this time using only the middle-skilled group as the lower-skilled workers. The results are in Tables 6–9. The point estimates are mostly very similar and still within the plausible range. Most of the point estimates of elasticity are considerably higher than 1.6, with half of the specifications yielding point estimates above two. However, the 95% confidence intervals usually do not contain the values above the strong skill-bias threshold (Tables 7–9). Again, the conventionally computed elasticities are again considerably higher, and in most cases implausibly so (Tables 7–9).

Tables 10–13 repeat the estimation, this time using the low-skilled workers as the second group. These results are again broadly similar to our baseline; however, this time the point estimates are somewhat higher and the confidence interval in many specifications includes the value of elasticity required for the strong form of skill bias. In the case of GMM, even the point estimate lie well into the strong bias region. Given that the definitions of middle- and low-skilled groups in EU KLEMS are not always consistent across countries or even across time for a given country, we are cautions with putting too much emphasize on the above results, preferring our baseline specification. However, it is reassuring that estimating our model for the two groups separately yields consistently similar findings.

https://doi.org/10.1017/S1365100522000372 Published online by Cambridge University Press

Table 6. High- vs. middle-skilled; OLS

	1	2	3	4	5	6
log (H/L)	-0.190^{***}	-0.190^{***}	-0.212^{**}	-0.207^{**}	-0.224^{***}	-0.181^{***}
	(0.068)	(0.068)	(0.088)	(0.091)	(0.068)	(0.066)
Country effects	No	No	No	No	No	No
Trend	Yes	Yes	Yes	Yes	Yes	No
Trend squared	No	Yes	No	Yes	Yes	No
Country trend	No	No	Yes	Yes	Yes	No
Country trend Sq	No	No	No	No	Yes	No
Year effects	No	No	No	No	No	Yes
σ	2.11	2.11	2.07	2.08	2.05	2.13
s.e.	(0.127)	(0.128)	(0.161)	(0.169)	(0.124)	(0.125)
95% conf. int.	[1.86, 2.36]	[1.85, 2.36]	[1.75, 2.39]	[1.74, 2.42]	[1.80, 2.30]	[1.88, 2.38]
p($\sigma < 2.5$)	1.00	1.00	0.99	0.98	1.00	1.00
σ'	5.26	5.26	4.72	4.83	4.46	5.54
R^2	0.29	0.29	0.53	0.53	0.78	0.29
N	642	642	642	642	642	642

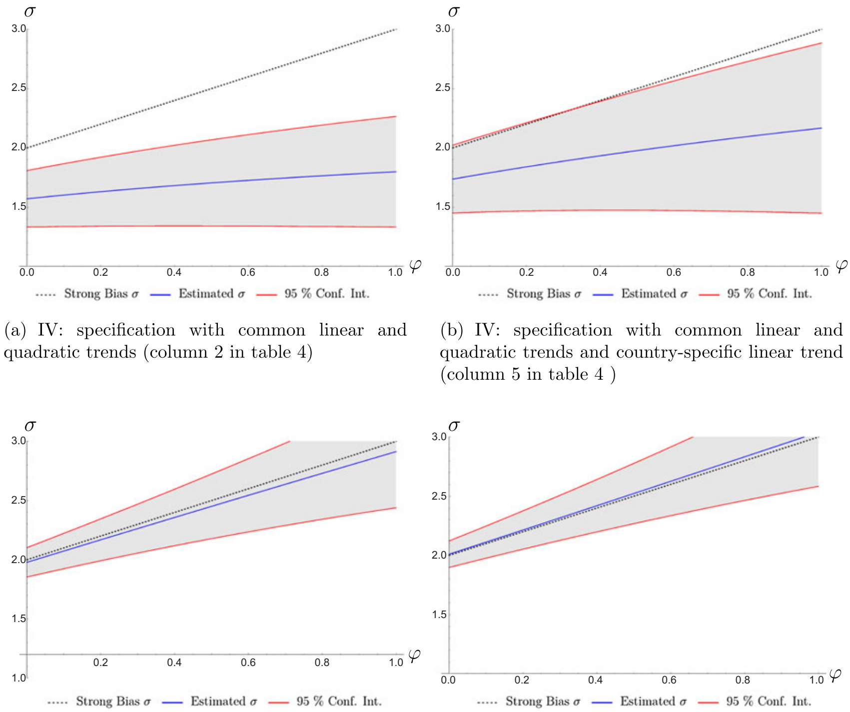
NOTES: Ordinary least squares estimates of $\log{(w_H/w_L)}_{it} = \gamma_0 + \gamma_1\,\mu(t) + \gamma_2\log{(H/L)}_{it} + \varepsilon_{it}$, using high-skilled workers as H and low-skilled workers for L. Standard errors clustered at the country level. "Macro-"elasticity of substitution σ calculated based on $\gamma_2 = (\sigma - 2 - \varphi)/(1 + \sigma\varphi)$ with $\varphi = 0.5$. Standard error of σ calculated using the Delta method. $p(\sigma < 2.5)$ denotes the probability that the true "macro-"elasticity is less than the strong skill-bias threshold. Conventional elasticity σ' calculated as $1/\gamma_2$.

Table 7. High- vs. middle-skilled; Country Fixed Effects

	1	2	3	4	5	6
log (H/L)	-0.468^{***}	-0.465^{***}	-0.528^{***}	-0.555^{***}	-0.621^{***}	-0.448^{***}
	(0.078)	(0.081)	(0.087)	(0.089)	(0.185)	(0.079)
Country effects	Yes	Yes	Yes	Yes	Yes	Yes
Trend	Yes	Yes	Yes	Yes	Yes	No
Trend squared	No	Yes	No	Yes	Yes	No
Country trend	No	No	Yes	Yes	Yes	No
Country trend Sq	No	No	No	No	Yes	No
Year effects	No	No	No	No	No	Yes
σ	1.65	1.65	1.56	1.52	1.43	1.68
s.e.	(0.115)	(0.120)	(0.122)	(0.122)	(0.243)	(0.118)
95% conf. int.	[1.42, 1.88]	[1.41, 1.89]	[1.32, 1.80]	[1.28, 1.77]	[0.95, 1.92]	[1.44, 1.91]
p($\sigma < 2.5$)	1.00	1.00	1.00	1.00	1.00	1.00
σ'	2.14	2.15	1.89	1.80	1.61	2.23
R^2	0.60	0.60	0.81	0.82	0.86	0.62
N	642	642	642	642	642	642

NOTES: Estimates of $\log{(w_H/w_L)}_{it} = \lambda_i + \gamma_0 + \gamma_1\,\mu(t) + \gamma_2\log{(H/L)}_{it} + \varepsilon_{it}$, using high-skilled workers as H and middle-skilled workers for L. Lagged values of relative labor supplies as instruments for the current level. Standard errors clustered at the country level. "Macro-"elasticity of substitution σ calculated based on $\gamma_2 = (\sigma - 2 - \varphi)/(1 + \sigma\varphi)$ with $\varphi = 0.5$. Standard error of σ calculated using the Delta method. $p(\sigma < 2.5)$ denotes the probability that the true "macro-"elasticity is less than the strong skill-bias threshold. Conventional elasticity σ' calculated as $1/\gamma_2$.

https://doi.org/10.1017/S1365100522000372 Published online by Cambridge University Press



(a) IV: specification with common linear and quadratic trends (column 2 in table 4)

(b) IV: specification with common linear and quadratic trends and country-specific linear trend (column 5 in table 4)

(c) GMM: specification with common linear and quadratic trends (column 2 in table 5)

(d) GMM: specification with common linear and quadratic trends and country-specific linear trend (column 5 in table 5)

Figure 2. Estimates of σ (blue line) and the two standard deviation intervals (shaded) for different values of the diffusion parameter φ based on different specifications under either IV or system GMM. The black line denotes the value required for strong skill bias.

4.2.3. Restricted sample

Finally, as explained above, we want to check if our results are not affected by changing definitions of skill categories. To this end, we re-run our regressions using only the original release of EU KLEMS data, thereby significantly reducing the possibility of changing skill definitions. Even for unchanging definition, there is a concern about changes in the amount of human capital supplied by workers with the same nominal education attainment but of different vintages which we have alluded to previously (Bowlus et al. 2017). This shorter sample also helps to further alleviate this concern. The results are presented in Tables 14–17. They are generally quite similar to our baseline results.

4.3 Sectoral analysis

The EU KLEMS data come disaggregated at the industry level according to the NACE European standard classification of productive economic activities. However, depending on the country and the time period, the level of disaggregation varies. Additionally, for later years in our sample (unfortunately the cutoff is not uniform across countries) the classification switches from NACE Revision1 to NACE revision 2, which split and combine some major sectors. For all these reason, a detailed sector-level analysis is beyond the scope of this paper as it would require a very detailed and careful merging of data from different time periods and countries. With those caveats in mind, we present the results of estimating our model for 9 major sectors separately: finance, health and

https://doi.org/10.1017/S1365100522000372 Published online by Cambridge University Press

Table 8. High- vs. middle-skilled; IV

	1	2	3	4	5	6
log (H/L)	-0.371^{***}	-0.359^{***}	-0.347^{***}	-0.320^{***}	-0.282^{***}	-0.334^{***}
	(0.066)	(0.070)	(0.056)	(0.085)	(0.058)	(0.065)
Country effects	Yes	Yes	Yes	Yes	Yes	Yes
Trend	Yes	Yes	Yes	Yes	Yes	No
Trend squared	No	Yes	No	Yes	Yes	No
Country trend	No	No	Yes	Yes	Yes	No
Country trend Sq	No	No	No	No	Yes	No
Year effects	No	No	No	No	No	Yes
σ	1.80	1.82	1.83	1.88	1.94	1.86
s.e.	(0.106)	(0.113)	(0.091)	(0.143)	(0.100)	(0.107)
95% conf. int.	[1.58, 2.01]	[1.59, 2.04]	[1.65, 2.02]	[1.59, 2.16]	[1.74, 2.14]	[1.64, 2.07]
p(σ < 2.5)	0.94	0.94	0.95	0.92	0.94	0.94
σ'	2.70	2.79	2.88	3.12	3.54	2.99
N	534	534	534	534	534	534

NOTES: IV estimates of $\log{(w_H/w_L)}_{it} = \lambda_i + \gamma_0 + \gamma_1\,\mu(t) + \gamma_2 \log{(H/L)}_{it} + \varepsilon_{it}$, using high-skilled workers as H and middle-skilled workers for L. Lagged values of relative labor supplies as instruments for the current level. Standard errors clustered at the country level. "Macro-"elasticity of substitution σ calculated based on $\gamma_2 = (\sigma - 2 - \varphi)/(1 + \sigma\varphi)$ with $\varphi = 0.5$. Standard error of σ calculated using the Delta method. $p(\sigma < 2.5)$ denotes the probability that the true "macro-"elasticity is less than the strong skill-bias threshold. Conventional elasticity σ' calculated as $1/\gamma_2$.

Table 9. High- vs. middle-skilled; system GMM

	1	2	3	4	5	6
log (H/L)	-0.207^{***}	-0.206^{***}	-0.205^{**}	-0.203^{**}	-0.225^{***}	-0.195^{***}
	(0.069)	(0.069)	(0.092)	(0.094)	(0.068)	(0.067)
Country effects	Yes	Yes	Yes	Yes	Yes	Yes
Trend	Yes	Yes	Yes	Yes	Yes	No
Trend squared	No	Yes	No	Yes	Yes	No
Country trend	No	No	Yes	Yes	Yes	No
Country trend Sq	No	No	No	No	Yes	No
Year effects	No	No	No	No	No	Yes
σ	2.08	2.08	2.08	2.09	2.04	2.10
s.e.	(0.127)	(0.128)	(0.170)	(0.174)	(0.123)	(0.125)
95% conf. int.	[1.83, 2.33]	[1.82, 2.33]	[1.74, 2.42]	[1.74, 2.43]	[1.80, 2.29]	[1.85, 2.35]
p(σ < 2.5)	1.00	1.00	0.98	0.98	1.00	1.00
σ'	4.84	4.85	4.88	4.92	4.44	5.12
N	642	642	642	642	642	642

NOTES: System GMM estimates of $\log{(w_H/w_L)}_{it} = \lambda_i + \gamma_0 + \gamma_1\,\mu(t) + \gamma_2 \log{(H/L)}_{it} + \varepsilon_{it}$, using high-skilled workers as H and low- and middle-skilled workers for L. Lagged levels of relative labor supplies as instruments for the first-difference equation and lagged differences of relative supplies are used as instruments in the level equation. Standard errors clustered at the country level. "Macro-"elasticity of substitution σ calculated based on $\gamma_2 = (\sigma - 2 - \varphi)/(1 + \sigma\varphi)$ with $\varphi = 0.5$. Standard error of σ calculated using the Delta method. $p(\sigma < 2.5)$ denotes the probability that the true "macro-"elasticity is less than the strong skill-bias threshold. Conventional elasticity σ' calculated as $1/\gamma_2$.

social work, transportation, mining and forestry, education, agriculture, manufacturing, construction, and wholesale trade. These results are presented in Table 18. For each estimation method, we present two specifications labeled "estimation method" 1 and "estimation method" 2 (e.g. OLS1 and OLS2), which correspond to columns 3 and 6 in our aggregate results tables; the first one includes a linear trend and a country-specific trend, the second includes year fixed effects. The

https://doi.org/10.1017/S1365100522000372 Published online by Cambridge University Press

Table 10. High- vs. low-skilled; OLS

	1	2	3	4	5	6
log (H/L)	0.019	0.018	0.043	0.043	0.030	0.015
	(0.048)	(0.048)	(0.038)	(0.038)	(0.043)	(0.050)
Country effects	No	No	No	No	No	No
Trend	Yes	Yes	Yes	Yes	Yes	No
Trend squared	No	Yes	No	Yes	Yes	No
Country trend	No	No	Yes	Yes	Yes	No
Country trend Sq	No	No	No	No	Yes	No
Year effects	No	No	No	No	No	Yes
σ	2.54	2.54	2.60	2.60	2.57	2.53
s.e.	(0.110)	(0.111)	(0.090)	(0.090)	(0.100)	(0.115)
95% conf. int.	[2.32, 2.76]	[2.32, 2.76]	[2.42, 2.78]	[2.42, 2.78]	[2.37, 2.77]	[2.30, 2.76]
p($\sigma < 2.5$)	0.35	0.36	0.13	0.14	0.25	0.39
σ'	−52.80	−55.46	−23.00	−23.30	−33.20	−67.78
R^2	0.03	0.03	0.68	0.70	0.86	−0.02
N	645	645	645	645	645	645

NOTES: Ordinary least squares estimates of $\log(w_H/w_L)_{it} = \gamma_0 + \gamma_1\,\mu(t) + \gamma_2 \log(H/L)_{it} + \varepsilon_{it}$, using high-skilled workers as H and middle-skilled workers for L on the restricted sample (1970-2005). Standard errors clustered at the country level. "Macro-"elasticity of substitution σ calculated based on $\gamma_2 = (\sigma - 2 - \varphi)/(1 + \sigma\varphi)$ with $\varphi = 0.5$. Standard error of σ calculated using the Delta method. $p(\sigma < 2.5)$ denotes the probability that the true "macro-"elasticity is less than the strong skill-bias threshold. Conventional elasticity σ' calculated as $1/\gamma_2$.

Table 11. High- vs. low-skilled; Country fixed effects

	1	2	3	4	5	6
log (H/L)	−0.153*	−0.160*	−0.068	−0.070	−0.142*	−0.182*
	(0.087)	(0.086)	(0.097)	(0.086)	(0.072)	(0.093)
Country effects	Yes	Yes	Yes	Yes	Yes	Yes
Trend	Yes	Yes	Yes	Yes	Yes	No
Trend squared	No	Yes	No	Yes	Yes	No
Country trend	No	No	Yes	Yes	Yes	No
Country trend Sq	No	No	No	No	Yes	No
Year effects	No	No	No	No	No	Yes
σ	2.18	2.17	2.35	2.35	2.20	2.12
s.e.	(0.168)	(0.166)	(0.204)	(0.180)	(0.142)	(0.176)
95% conf. int.	[1.84, 2.52]	[1.84, 2.50]	[1.94, 2.76]	[1.99, 2.71]	[1.92, 2.49]	[1.77, 2.48]
p($\sigma < 2.5$)	0.96	0.96	0.76	0.79	0.97	0.97
σ'	6.52	6.24	14.63	14.35	7.06	5.49
R^2	0.07	0.08	0.71	0.71	0.82	0.10
N	645	645	645	645	645	645

NOTES: Estimates of $\log(w_H/w_L)_{it} = \lambda_i + \gamma_0 + \gamma_1\,\mu(t) + \gamma_2 \log(H/L)_{it} + \varepsilon_{it}$, using high-skilled workers as H and low-skilled workers for L. Standard errors clustered at the country level. "Macro-"elasticity of substitution σ calculated based on $\gamma_2 = (\sigma - 2 - \varphi)/(1 + \sigma\varphi)$ with $\varphi = 0.5$. Standard error of σ calculated using the Delta method. $p(\sigma < 2.5)$ denotes the probability that the true "macro-"elasticity is less than the strong skill-bias threshold. Conventional elasticity σ' calculated as $1/\gamma_2$.

Table 12. High- vs. low-skilled; IV

	1	2	3	4	5	6
log (H/L)	−0.156*	−0.180*	−0.016	−0.025	−0.207***	−0.189*
	(0.095)	(0.092)	(0.108)	(0.098)	(0.043)	(0.097)
Country effects	Yes	Yes	Yes	Yes	Yes	Yes
Trend	Yes	Yes	Yes	Yes	Yes	No
Trend squared	No	Yes	No	Yes	Yes	No
Country trend	No	No	Yes	Yes	Yes	No
Country trend Sq	No	No	No	No	Yes	No
Year effects	No	No	No	No	No	Yes
σ	2.17	2.13	2.47	2.45	2.08	2.11
s.e.	(0.183)	(0.174)	(0.240)	(0.214)	(0.079)	(0.182)
95% conf. int.	[1.81, 2.54]	[1.78, 2.48]	[1.99, 2.94]	[2.02, 2.87]	[1.92, 2.24]	[1.75, 2.47]
p($\sigma < 2.5$)	0.83	0.85	0.55	0.58	0.94	0.85
σ'	6.41	5.57	63.96	40.54	4.82	5.29
N	537	537	537	537	537	537

NOTES: IV estimates of $\log (w_H/w_L)_{it} = \lambda_i + \gamma_0 + \gamma_1\,\mu(t) + \gamma_2 \log (H/L)_{it} + \varepsilon_{it}$, using high-skilled workers as H and middle-skilled workers for L. Lagged values of relative labor supplies as instruments for the current level. Standard errors clustered at the country level. "Macro-"elasticity of substitution σ calculated based on $\gamma_2 = (\sigma - 2 - \varphi)/(1 + \sigma\varphi)$ with $\varphi = 0.5$. Standard error of σ calculated using the Delta method. $p(\sigma < 2.5)$ denotes the probability that the true "macro-"elasticity is less than the strong skill-bias threshold. Conventional elasticity σ' calculated as $1/\gamma_2$.

Table 13. High- vs. low-skilled; System GMM

	1	2	3	4	5	6
log (H/L)	0.015	0.013	0.053	0.052	0.025	0.010
	(0.052)	(0.052)	(0.042)	(0.042)	(0.043)	(0.054)
Country effects	Yes	Yes	Yes	Yes	Yes	Yes
Trend	Yes	Yes	Yes	Yes	Yes	No
Trend squared	No	Yes	No	Yes	Yes	No
Country trend	No	No	Yes	Yes	Yes	No
Country trend Sq	No	No	No	No	Yes	No
Year effects	No	No	No	No	No	Yes
σ	2.53	2.53	2.62	2.62	2.56	2.52
s.e.	(0.118)	(0.119)	(0.099)	(0.099)	(0.100)	(0.123)
95% conf. int.	[2.30, 2.77]	[2.29, 2.77]	[2.42, 2.82]	[2.42, 2.82]	[2.36, 2.76]	[2.28, 2.77]
p($\sigma < 2.5$)	0.39	0.40	0.11	0.11	0.28	0.43
σ'	−68.39	−74.80	−18.95	−19.28	−39.99	−99.24
N	645	645	645	645	645	645

NOTES: System GMM estimates of $\log (w_H/w_L)_{it} = \lambda_i + \gamma_0 + \gamma_1\,\mu(t) + \gamma_2 \log (H/L)_{it} + \varepsilon_{it}$, using high-skilled workers as H and low- and low-skilled workers for L. Lagged levels of relative labor supplies as instruments for the first-difference equation and lagged differences of relative supplies are used as instruments in the level equation. Standard errors clustered at the country level. "Macro-"elasticity of substitution σ calculated based on $\gamma_2 = (\sigma - 2 - \varphi)/(1 + \sigma\varphi)$ with $\varphi = 0.5$. Standard error of σ calculated using the Delta method. $p(\sigma < 2.5)$ denotes the probability that the true "macro-"elasticity is less than the strong skill-bias threshold. Conventional elasticity σ' calculated as $1/\gamma_2$.

fixed effects, IV, and GMM methods all include country fixed effects, as before. The findings are similar to the aggregate results and well within the plausible range (unlike the conventional estimates) but in many cases have slightly higher values. Interestingly, in some instances, most notably in the finance and the education sectors, there is quite pronounced indication of a strong skill bias.

https://doi.org/10.1017/S1365100522000372 Published online by Cambridge University Press

Table 14. High- vs. lower-skilled ; OLS (1970–2005)

	1	2	3	4	5	6
log (H/L)	−0.014	−0.018	0.010	0.006	0.015	−0.019
	(0.055)	(0.054)	(0.052)	(0.052)	(0.061)	(0.057)
Country effects	No	No	No	No	No	No
Trend	Yes	Yes	Yes	Yes	Yes	No
Trend squared	No	Yes	No	Yes	Yes	No
Country trend	No	No	Yes	Yes	Yes	No
Country trend Sq	No	No	No	No	Yes	No
Year effects	No	No	No	No	No	Yes
σ	2.47	2.46	2.52	2.51	2.54	2.46
s.e.	(0.123)	(0.120)	(0.119)	(0.118)	(0.140)	(0.125)
95% conf. int.	[2.22, 2.71]	[2.22, 2.70]	[2.29, 2.76]	[2.28, 2.75]	[2.26, 2.81]	[2.21, 2.71]
p($\sigma < 2.5$)	0.60	0.63	0.42	0.45	0.40	0.63
σ'	71.46	55.78	−97.94	−155.79	−64.75	54.03
R^2	0.02	0.06	0.53	0.54	0.76	0.00
N	413	413	413	413	413	413

NOTES: Ordinary least squares estimates of $\log(w_H/w_L)_{it} = \gamma_0 + \gamma_1\,\mu(t) + \gamma_2 \log(H/L)_{it} + \varepsilon_{it}$, using high-skilled workers as H and middle-skilled workers for L on the restricted sample (1970-2005). Standard errors clustered at the country level. "Macro-"elasticity of substitution σ calculated based on $\gamma_2 = (\sigma - 2 - \varphi)/(1 + \sigma\varphi)$ with $\varphi = 0.5$. Standard error of σ calculated using the Delta method. $p(\sigma < 2.5)$ denotes the probability that the true "macro-"elasticity is less than the strong skill-bias threshold. Conventional elasticity σ' calculated as $1/\gamma_2$.

Table 15. High- vs. lower-skilled; Country fixed effects (1970–2005)

	1	2	3	4	5	6
log (H/L)	-0.367^{**}	-0.363^{**}	-0.225^{**}	-0.240^{***}	-0.298^{*}	-0.369^{**}
	(0.138)	(0.133)	(0.077)	(0.067)	(0.141)	(0.135)
Country effects	Yes	Yes	Yes	Yes	Yes	Yes
Trend	Yes	Yes	Yes	Yes	Yes	No
Trend squared	No	Yes	No	Yes	Yes	No
Country trend	No	No	Yes	Yes	Yes	No
Country trend Sq	No	No	No	No	Yes	No
Year effects	No	No	No	No	No	Yes
σ	1.80	1.81	2.05	2.02	1.92	1.80
s.e.	(0.222)	(0.215)	(0.140)	(0.120)	(0.241)	(0.216)
95% conf. int.	[1.36, 2.25]	[1.38, 2.24]	[1.77, 2.32]	[1.78, 2.26]	[1.43, 2.40]	[1.37, 2.23]
p($\sigma < 2.5$)	0.99	0.99	0.99	1.00	0.97	0.99
σ'	2.73	2.76	4.45	4.17	3.35	2.71
R^2	0.20	0.22	0.71	0.72	0.84	0.17
n	413	413	413	413	413	413

NOTES: Estimates of $\log(w_H/w_L)_{it} = \lambda_i + \gamma_0 + \gamma_1\,\mu(t) + \gamma_2 \log(H/L)_{it} + \varepsilon_{it}$, using high-skilled workers as H and middle-skilled workers for L on the restricted sample (1970-2005). Standard errors clustered at the country level. "Macro-"elasticity of substitution σ calculated based on $\gamma_2 = (\sigma - 2 - \varphi)/(1 + \sigma\varphi)$ with $\varphi = 0.5$. Standard error of σ calculated using the Delta method. $p(\sigma < 2.5)$ denotes the probability that the true "macro-"elasticity is less than the strong skill-bias threshold. Conventional elasticity σ' calculated as $1/\gamma_2$.

Table 16. High vs. lower-skilled ; IV (1970–2005)

	1	2	3	4	5	6
log (H/L)	-0.423^{***}	-0.419^{***}	-0.120	-0.130^{*}	-0.281^{***}	-0.417^{***}
	(0.139)	(0.137)	(0.083)	(0.076)	(0.093)	(0.140)
Country effects	Yes	Yes	Yes	Yes	Yes	Yes
Trend	Yes	Yes	Yes	Yes	Yes	No
Trend squared	No	Yes	No	Yes	Yes	No
Country trend	No	No	Yes	Yes	Yes	No
Country trend Sq	No	No	No	No	Yes	No
Year effects	No	No	No	No	No	Yes
σ	1.71	1.72	2.25	2.23	1.95	1.72
s.e.	(0.214)	(0.211)	(0.166)	(0.152)	(0.160)	(0.216)
95% conf. int.	[1.29, 2.14]	[1.30, 2.14]	[1.91, 2.58]	[1.92, 2.53]	[1.62, 2.27]	[1.29, 2.15]
p(σ < 2.5)	0.90	0.90	0.81	0.83	0.90	0.90
σ'	2.37	2.39	8.35	7.70	3.56	2.40
N	368	368	368	368	368	368

NOTES: IV estimates of $\log (w_H/w_L)_{it} = \lambda_i + \gamma_0 + \gamma_1\, \mu(t) + \gamma_2 \log (H/L)_{it} + \varepsilon_{it}$, using high-skilled workers as H and low- and middle-skilled workers for L on the restricted sample (1970–2005). Lagged values of relative labor supplies as instruments for the current level. Standard errors clustered at the country level. "Macro-"elasticity of substitution σ calculated based on $\gamma_2 = (\sigma - 2 - \varphi)/(1 + \sigma\varphi)$ with $\varphi = 0.5$. Standard error of σ calculated using the Delta method. $p(\sigma < 2.5)$ denotes the probability that the true "macro-"elasticity is less than the strong skill-bias threshold. Conventional elasticity σ' calculated as $1/\gamma_2$.

Table 17. High vs. lower-skilled; System GMM (1970–2005)

	1	2	3	4	5	6
log (H/L)	-0.014	-0.018	0.011	0.007	0.014	-0.018
	(0.055)	(0.054)	(0.053)	(0.053)	(0.061)	(0.057)
Country effects	Yes	Yes	Yes	Yes	Yes	Yes
Trend	Yes	Yes	Yes	Yes	Yes	No
Trend squared	No	Yes	No	Yes	Yes	No
Country trend	No	No	Yes	Yes	Yes	No
Country trend Sq	No	No	No	No	Yes	No
Year effects	No	No	No	No	No	Yes
σ	2.47	2.46	2.52	2.52	2.53	2.46
s.e.	(0.123)	(0.120)	(0.120)	(0.119)	(0.140)	(0.125)
95% conf. int.	[2.22, 2.71]	[2.22, 2.70]	[2.29, 2.76]	[2.28, 2.75]	[2.25, 2.81]	[2.21, 2.71]
p(σ < 2.5)	0.60	0.63	0.42	0.45	0.41	0.62
σ'	71.00	56.37	-93.97	-144.89	-69.41	54.39
N	413	413	413	413	413	413

NOTES: System GMM estimates of $\log (w_H/w_L)_{it} = \lambda_i + \gamma_0 + \gamma_1\, \mu(t) + \gamma_2 \log (H/L)_{it} + \varepsilon_{it}$, using high-skilled workers as H and low- and middle-skilled workers for L. Lagged levels of relative labor supplies as instruments for the first-difference equation and lagged differences of relative supplies are used as instruments in the level equation. Standard errors clustered at the country level. "Macro-"elasticity of substitution σ calculated based on $\gamma_2 = (\sigma - 2 - \varphi)/(1 + \sigma\varphi)$ with $\varphi = 0.5$. Standard error of σ calculated using the Delta method. $p(\sigma < 2.5)$ denotes the probability that the true "macro-"elasticity is less than the strong skill-bias threshold. Conventional elasticity σ' calculated as $1/\gamma_2$.

https://doi.org/10.1017/S1365100522000372 Published online by Cambridge University Press

Table 18. High- vs. lower-skilled; Sectoral regressions

	OLS 1	OLS 2	FE 1	FE 2	IV 1	IV 2	GMM 1	GMM 2
Aggregate								
σ	2.46	2.51	1.89	1.85	1.96	1.75	2.52	2.45
s.e.	(0.140)	(0.110)	(0.137)	(0.188)	(0.156)	(0.194)	(0.129)	(0.135)
$p(\sigma < 2.5)$	0.62	0.47	1.00	1.00	0.90	0.90	0.44	0.63
Finance								
σ	2.81	2.92	2.25	2.82	2.26	2.96	2.81	2.82
s.e.	(0.201)	(0.153)	(0.195)	(0.272)	(0.448)	(0.310)	(0.220)	(0.172)
$p(\sigma < 2.5)$	0.05	0.00	0.88	0.11	0.65	0.17	0.07	0.03
Health & social								
σ	2.64	2.68	2.18	1.93	2.16	1.84	2.65	2.68
s.e.	(0.246)	(0.206)	(0.116)	(0.153)	(0.123)	(0.182)	(0.247)	(0.233)
$p(\sigma < 2.5)$	0.29	0.18	0.99	1.00	0.88	0.90	0.27	0.22
Transportation								
σ	2.23	2.42	1.70	1.92	2.21	2.24	2.21	2.31
s.e.	(0.250)	(0.209)	(0.238)	(0.087)	(0.300)	(0.131)	(0.251)	(0.222)
$p(\sigma < 2.5)$	0.84	0.65	1.00	1.00	0.74	0.84	0.86	0.79
Mining								
σ	2.35	2.43	1.82	1.90	1.94	2.08	2.32	2.37
s.e.	(0.234)	(0.172)	(0.344)	(0.245)	(0.318)	(0.226)	(0.241)	(0.211)
$p(\sigma < 2.5)$	0.73	0.66	0.95	0.98	0.82	0.83	0.76	0.72
Education								
σ	2.73	2.81	2.71	2.83	2.78	2.82	2.72	2.74
s.e.	(0.087)	(0.086)	(0.171)	(0.158)	(0.207)	(0.161)	(0.083)	(0.082)
$p(\sigma < 2.5)$	0.00	0.00	0.11	0.02	0.19	0.14	0.00	0.00
Agriculture								
σ	2.46	2.57	2.44	2.43	2.55	2.43	2.48	2.51
s.e.	(0.185)	(0.169)	(0.208)	(0.229)	(0.239)	(0.249)	(0.192)	(0.183)
$p(\sigma < 2.5)$	0.59	0.33	0.62	0.62	0.44	0.58	0.55	0.49
Manufacturing								
σ	2.54	2.54	2.19	1.96	2.20	1.92	2.57	2.49
s.e.	(0.136)	(0.104)	(0.238)	(0.193)	(0.408)	(0.231)	(0.141)	(0.107)
$p(\sigma < 2.5)$	0.40	0.37	0.88	0.99	0.69	0.86	0.32	0.53
Construction								
σ	2.47	2.52	2.24	2.35	2.13	2.38	2.48	2.51
s.e.	(0.038)	(0.053)	(0.154)	(0.194)	(0.207)	(0.296)	(0.037)	(0.054)
$p(\sigma < 2.5)$	0.75	0.33	0.94	0.77	0.83	0.62	0.71	0.41
Wholesale								
σ	2.68	2.71	2.02	2.29	2.12	2.36	2.67	2.66
s.e.	(0.115)	(0.108)	(0.257)	(0.177)	(0.322)	(0.234)	(0.123)	(0.117)
$p(\sigma < 2.5)$	0.06	0.03	0.95	0.87	0.76	0.67	0.08	0.09

NOTES: Estimates using various specifications and methods at sectoral level. For each method, specification 1 includes (a) a common trend and (b) a country-specific trend, and specification (2) includes year effects. All regressions using high-skilled workers as H and low- and middle-skilled workers for L. Standard errors clustered at the country level. All other details as in the previous tables.

https://doi.org/10.1017/S1365100522000372 Published online by Cambridge University Press

The sectoral results provide a good context to mention an important point, namely that the use of education as a basis for classifying workers may be a limiting factor in our analysis.[9] Education is very likely too broad a concept to fully capture the nature of tasks performed by employees in a modern economy. Often work characteristics, such as cognitive requirements, routine vs. creative, reliance on teams and cooperation, etc. differed significantly across jobs that are performed by workers with the same nominal education level. Unfortunately, the EU KLEMS labor data we are working only contain three skill categories and are devoid of useful information on occupation, position in firm hierarchy, tenure with firm, age or experience, etc. The closest, but obviously crude and incomplete, way we can come getting at this issue with our data is by thinking about sectorial analyst. Perhaps, workers with similar education levels perform different tasks in different sectors/industries, rendering them more or less substitutable for workers with a different education depending on the nature of those tasks (and thus the industry of employment). However, to fully explore this issue one would have to extend our theoretical approach to include a model with differential tasks (Acemoglu and Autor, 2011; Acemoglu and Restrepo, 2019) and it would likely require data sources beyond what is available in EU KLEMS.

5. Conclusions

Aggregate elasticity of substitution between workers with different skill levels is an important macroeconomics parameter. For example, it is crucial for quantifying the impact of technological and structural changes on the macroeconomy since it determines how shifts in labor composition and technology affect relative wages (Acemoglu, 1998, 2002; Acemoglu and Autor, 2011; Krusell et al. 2000). Importantly, with sufficiently high degree of substitutability between skills, endogenous-directed technological change can lead to a *strong skill-bias*: a situation where an increase in the relative supply of skilled workers can—counter to the standard negative supply effect—raise the wages of those workers, helping to explain the secular rise in the skill premium over the last several decades. Additionally, the elasticity of substitution between skill types has played a crucial role in quantitative modeling of international income differences, the so-called *development accounting*, with the relative importance of human capital endowment usually hinging on its value (Caselli and Coleman, 2006; Jerzmanowski and Tamura, 2019).

The empirical work seeking to estimate the value of elasticity of substitution has, largely relaying on US microdata, led to a consensus value of about 1.6, which is not high enough for the strong skill-bias to occur. In the economic growth literature, this numerical value of the elasticity—when used in development accounting studies—usually implies a large role of human capital in explaining cross-country income differences. However, there are reasons to be cautious about the existing estimates. First, it is not clear whether using an elasticity value obtained from US microdata is suitable for calibrating models aimed at explaining the behavior of widely diverse groups of economies (Jones, 2014). Second, the assumption of constancy of human capital quality over extended time periods—implied when using long US micro-time series to estimate the elasticity—may not be justified (Bowlus et al. 2017). We contribute to this literature by estimating the elasticity of substitution between workers of different skill types using a macropanel data from a *large group of economies* with most observations coming from only the more *recent time periods*.

Using an endogenous-directed technology model with international diffusion of ideas, we derive the appropriate estimating equation and show that—in a cross-country setting with technology diffusion—the elasticity is not a simple inverse of the wage/labor supply regression coefficient. We estimate this equation using data from the EU KLEMS Growth and Productivity Accounts panel data set, a detailed database of industry-level measures of output, inputs, and productivity for 28 European countries, Japan, South Korea, Australia and the USA for the period from 1970 to 2015. We find that the elasticity of substitution between tertiary-educated workers and those with lower education levels likely falls within the range of 1.7 and 2.5, which is higher than previous estimates but within a plausible range. Notably, our elasticity estimates are closer to those obtained recently by Bowlus et al., (2017), who argue that most of the past literature mis-measured changes in skill supplies over long time periods. Our approach does not explicitly correct for this problem but—given that our use of panel data allows us to get by with a shorter time dimensions of the sample—this should be less of a concern for us. As a result, the fact that our estimates also point to higher elasticity of substitution values is reassuring. In most of our regressions, the estimated elasticity falls short of the value required for strong skill-bias of technology; however, in some specifications, the 95% confidence interval includes that level, suggesting strong skill bias is possible. When we restrict our attention to a narrower definition of low-skilled workers (those form the low-skill group only), the findings point even more strongly in the direction of higher elasticity and a possibility of a strong skill-bias. Finally, we also estimate our model using disaggregated data for nine major sectors. They are similar to the aggregate results and well within the plausible range but in many cases have slightly higher values. Interestingly, in some instances, most notably in the finance and the education sectors, there is quite pronounced indication of a strong skill bias.

Our conclusions for this exercise are as follows. First, using a macropanel data of a fairly diverse group of countries over an extended time period, the elasticity estimates are largely in line with existing estimates obtained using very different data sets. We feel this is encouraging and strengthens the confidence we have in the range of plausible elasticity values. Second, we demonstrate that when using cross-country data, accounting for technology diffusion and endogeneity of the direction of technological change (what we dub the "macro-approach") are important for correctly interpreting the elasticity estimates. We feel that further work using international data to investigate elasticity of substitution between skills should continue to pay close attention to the underlying theoretical structure. Finally, the fact that some of our results point in the direction of higher elasticities is also consistent with the most recent findings using more disaggregated data, further strengthening their plausibility. A more robust way of addressing the potential role of endogeneity of the labor supply, accounting for off-balanced growth path dynamics, and incorporating data for less developed countries are all areas where our research could fruitfully be extended. As emphasized in the previous section, going beyond education differences and exploring the role of finer skills, tasks, and other job characteristics in making workers more or less substitutable is also of great interest and importance.[10]

Notes

1 Hendricks and Schoellman (2020) are another recent contribution which seeks to quantify international income differences using an approach based on imperfect substitutability of workers with different skill levels.

2 Howevet, the addition of more flexible parameterization of the technology term brings that estimate down to about 1.8.

3 In the appendix, we also provide a linearization of the model, which can serve as a basis for non-linear estimation of the elasticity without the assumption of balanced growth path. We leave this extension for future work.

4 We also run specifications with country fixed effects since η_H/η_L could be country-specific, as in Tamura, et al. (2016), and Tamura (2006).

5 Appendix C provides a detailed illustration of the comparison between this macrointerpretation and the conventional approach.

6 The results for only high and middle-skilled workers are provided in the appendix. They are similar to our main results, but the elasticity estimates are somewhat lower.

7 We do not use the data after 2015 since the source information for wage shares after 2015 is not available, and that series in the most recent release is simply interpolated from older data (Adarov and Stehrer, 2019). Releases up to 2016 can be found at http://www.euklems.net/ while the latest release is available at https://euklems.eu/.

8 Arellano and Bond (1991) derive a GMM estimator that uses suitably lagged levels of the dependent and predetermined right-hand side variables as instruments for the equation in first differences. Blundell and Bond (1998) extend this method to a system GMM estimator, where sufficiently lagged first differences are used to instrument an additional equation in levels, and this is the GMM estimator we use.

9 We thank one of the referees for bringing this to our attention
10 This set-up leads to an aggregate production function with unitary elasticity of substitution between capital and skilled labor. Evidence using cross-country aggregate data suggest this elasticity is less than one (e.g. Jerzmanowski, 2007) but we abstract from this issue here.

References

Acemoglu, D. (1998) Why do new technologies complement skills? Directed technical change and wage inequality. *Quarterly Journal of Economics* 113(4), 1055–1090.

Acemoglu, D. (2002) Directed technical change. *Review of Economic Studies* 69(4), 781–809.

Acemoglu, D. (2009) *Introduction to Modern Economic Growth*. Princeton, NJ: Princeton University Press.

Acemoglu, D. and D. Autor (2011) Skills, Tasks and Technologies: Implications for Employment and Earnings. **In:** Acemoglu, D. and D. Autor (eds.), *Handbook of Labor Economics*, vol. 4, pp. 1043–1171, Elsevier B.V.

Acemoglu, D. and P. Restrepo (2019) Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives* 3(2), 3–30.

Adarov, Amat and Robert Stehrer (2019) Tangible and Intangible Assets in the Growth Performance of the EU, Japan and the US. https://wiiw.ac.at/p-5058.html. [Online; accessed 5/5/20].

Arellano, M. and S. Bond (1991) Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies* 58, 277–297.

Arellano, M., and O. Bover (1995) Another look at the instrumental variable estimation of error-component models. *Journal of Econometrics* 68, 29–52.

Barro, R. J. (2012) Convergence and modernization revisited. Working paper 18295. *National Bureau of Economic Research*, 10.3386/w18295, http://www.nber.org/papers/w18295

Blundell, R. and S. Bond (1998) Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics* 87, 115–143.

Bowlus, A., E. Bozkurt, L. Lochner and C. Robinson (2017) Wages and employment: The canonical model revisited. Working paper 24069. *National Bureau of Economic Research*, 10.3386/w24069, http://www.nber.org/papers/w24069

Caselli, F. and C. Antonio (2019) The human capital stock: A generalized approach: Comment. *American Economic Review* 109.3, 1155–1174

Caselli, F. and W. J. Coleman (2006) The world technology frontier. *American Economic Review* 96(3), 499–522.

Ciccone, A. and G. Peri (2005) The Long-Run substitutability between more and less educated workers: Evidence from U.S. States, 1950–1990. *The Review of Economics and Statistics* 87(4), 652–663.

EU KLEMS project on Growth and Productivity in the European Union (2018) EU KLEMS Data. http://www.euklems.net/index.html. [Online; accessed 10/12/18].

Greene, W. H. (2003) Econometric Analysis, 5th edition, Pearson Education, isbn: 0-13-066189-9, http://pages.stern.nyu.edu/~wgreene/Text/econometricanalysis.htm.

Heckman, J., L. Lochner and C. Taber (1998) Explaining rising wage inequality: Explanations with a dynamic general equilibrium model of labor earnings with heterogeneous agents. *Review of Economic Dynamics* 1.1, 1–58. doi: 10.1006/redy.1997.0008, https://ideas.repec.org/a/red/issued/v1y1998i1p1-58.html

Hendricks, L. and T. Schoellman (2020) Skilled labor productivity and cross-country income differences. *Working Paper*, https://lhendricks.org/Research/jones_cc/paper_jcc.pdf

Jäger, K. (July 2018) Conference Board Report. http://www.euklems.net/index.html. [Online; accessed 12/18/2019].

Jerzmanowski, M. (2007) TFP differences: Appropriate technology vs. efficiency. *European Economic Review* 51(8), 1963–2110.

Jerzmanowski, M. and R. Tamura (2019) Directed technological change and cross-country income differences: A quantitative analysis. *Journal of Development Economics* 141, 10.1016/j.jdeveco.2019.102372, https://ideas.repec.org/a/eee/deveco/v141y2019ics0304387818305820.html

Johnson, G. (1970) The demand for labor by educational category. *Southern Economic Journal* 37(2), 190–204.

Jones, B. F. (2014) The human capital stock: A generalized approach. *American Economic Review* 104.11, 3752–3777. doi: 10.1257/aer.104.11.3752, http://www.aeaweb.org/articles?id10.1257/aer.104.11.3752

Katz, L. F. and K. M. Murphy (1992) Changes in relative wages, 1963-1987: Supply and demand factors. *The Quarterly Journal of Economics* 107.1, 35–78

Klenow, P. and A. Rodríguez-Clare (1997) The Neoclassical Revival in Growth Economics: Has It Gone Too Far?, *NBER Macroeconomics Annual 1997*, vol. 12, pp. 73–114, NBER Chapters. National Bureau of Economic Research, Inc, https://ideas.repec.org/h/nbr/nberch/11037.html

Krusell, P., L. E. Ohanian, J.-V. Rios-Rull and G. L. Violante (2000) Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica* 68(5), 1029–1054.

O'Mahony, M. and M. P. Timmer (June 2009) Output, input and productivity measures at the industry level: The EU KLEMS database. *The Economic Journal* 119(538), 374–403.

Tamura, R. (2006) Human capital and economic development. *Journal of Development Economics* 79(1), 26–72.

Tamura, R., C. J. Simon and K. M. Murphy (2016) Black and white fertility, differential baby booms: The value of equal education opportunity. *Journal of Demographic Economics* 82(1), 27–109.

Timmer, M., T. van Moergastel, E. Stuivenwold, G. Ypma, M. O'Mahony and M. Kangasniemi (2007) EU KLEMS Growth and Productivity Accounts Version 1.0; Part I Methodology. http://www.euklems.net/data/EUKLEMS_Growth_and_Productivity_Accounts_Part_I_Methodology.pdf. [Online; accessed 12/12/19].

Appendix A: Detailed derivations

In this section of the appendix, we provide more detailed derivations of some of the equations on the production side of the model.

A.1. Households

There is a continuum N of infinitely lived representative households with CRRA preferences, a discount rate of ρ and one of three skill types: high-skilled (H) and low-skilled (L). Population and type shares, s_i, are constant

The households own physical capital and patents rights on innovation and maximize the present discounted value of an infinite stream of utility. The optimal consumption path obeys the familiar Euler equation

$$\frac{\dot{C}}{C} = \frac{1}{\theta}\left[r - \rho\right]$$

where ρ is the discount rate, θ is the CRRA coefficient, and the interest rate r is equal to the rental rate minus the rate of depreciation.

A.2. Final good

Final output is produced using intermediate goods which are skill-specific according to the following production function

$$Y = \{Y_H^{\frac{\varepsilon-1}{\varepsilon}} + Y_L^{\frac{\varepsilon-1}{\varepsilon}}\}^{\frac{\varepsilon}{\varepsilon-1}}. \tag{10}$$

Competitive firms (characterized below) produce the intermediate goods Y_H and Y_L and sell them to competitive final output producers at prices P_i, $i = H, L$. We take the final good to be the numeraire so that

$$\left[P_L^{1-\varepsilon} + P_H^{1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}} = 1. \tag{11}$$

A.3. Intermediate goods & Machines

Intermediate goods producers combine machines and labor in the standard "variety of machine inputs" manner

$$Y_L = \frac{1}{1-\beta}\int_0^{A_L} \chi_{jL}^{1-\beta}\, dj\, L^{\beta} \tag{12}$$

where χ_{jL} is quantity of machines of variety j rented by the L-type intermediate goods producer.

The representative L-type intermediate goods firm solves the following maximization problem:

$$\max_{\{\chi_{jL}, L\}} \left\{\frac{P_L}{1-\beta}\int_0^{A_L} \chi_{jL}^{1-\beta}\, djL^{\beta} - \int_0^{A_L} p_{jL}\chi_{jL}dj - w_L L\right\}, \tag{13}$$

where p_{jL} is the price of variety j, L-type machine.

For a representative firm hiring workers of skill type L, the inverse derived demand for a typical machine j is given by

$$P_L \chi_{jL}^{-\beta} L^{\beta} = p_{jL} \tag{14}$$

Blueprints for machines varieties are specific to the economy. They are invented by local entrepreneurs who hold perpetual monopoly rights over a given variety they have invented within the country. Machines are supplied to the intermediate goods producers by the monopolists who own the blueprints and rent capital to manufacture the machines.[10] Capital is rented in a competitive market at the capital rental rate R. One unit of physical capital can produce one machine of any variety and machines depreciate at a rate of 100%. Each machine producing monopolist faces a potential imitator with cost $\upsilon > 1$ times higher the original innovator's own marginal cost, which implies that they will set the price equal to a υ markup over her own marginal cost.

$$p_{jL} = \upsilon R \tag{15}$$

The equilibrium supply of machines of type j to skill L, and the equilibrium quantities of machines are:

$$\chi_{jL} = \left(\frac{P_L}{\upsilon R}\right)^{1/\beta} L \tag{16}$$

which means the (derived) production functions of intermediate goods become

$$Y_L = \frac{1}{1-\beta}\left(\frac{P_L}{\upsilon R}\right)^{\frac{1-\beta}{\beta}} A_L L \tag{17}$$

And the profit per line of machines is given by

$$\pi_{jL} = \left(\frac{\upsilon - 1}{\upsilon}\right) P_L^{1/\beta} L \,(\upsilon R)^{\frac{\beta-1}{\beta}} \tag{18}$$

Finally, it also follows that the relative prices of the two intermediate goods are given by:

$$\frac{P_H}{P_L} = \left(\frac{A_H H}{A_L L}\right)^{-\frac{\beta}{\sigma}} \tag{19}$$

where $\sigma = 1 + (\varepsilon - 1)\beta$.

A.4. Wages & Technology

Intermediate goods producers hire labor according to the following first-order condition:

$$\frac{\beta P_L}{1-\beta}\int_0^{A_L} \chi_{jL}^{1-\beta} dj L^{\beta-1} = w_L, \tag{20}$$

which, after substituting for the equilibrium quantities of machines and available workers of type L, produces

$$w_L = \frac{\beta}{1-\beta} A_L \beta P_L^{\frac{1}{\beta}} (\upsilon R)^{-\frac{1-\beta}{\beta}} \tag{21}$$

Thus, the relative wages of workers with different skill levels are given by

$$\frac{w_H}{w_L} = \left(\frac{A_H}{A_L}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{H}{L}\right)^{-\frac{1}{\sigma}} \tag{22}$$

where σ is the elasticity of substitution between H and L.

https://doi.org/10.1017/S1365100522000372 Published online by Cambridge University Press

A.5. Capital allocation & rental rate

Capital is used to manufacture machines. Denoting by K_L the amount of physical capital devoted to production of L-type machines, we have

$$K_L = \int_0^{A_L} \chi_{jL} dj = A_L \left(\frac{P_L}{\upsilon R}\right)^{1/\beta} L \tag{23}$$

and it follows that

$$\frac{K_H}{K_L} = \left(\frac{A_H}{A_L}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{H}{L}\right)^{\frac{\sigma-1}{\sigma}} \tag{24}$$

In addition, total capital stock by K is given by

$$K = K_L + K_H. \tag{25}$$

Since machines take one unit of capital to produce and all machines within a skill industry are symmetric, it must be the case that:

$$K_L = A_L \chi_L \tag{26}$$

Substituting this into the intermediate goods production functions

$$Y_L = \frac{1}{1-\beta} K_L^{1-\beta} (A_L L)^{\beta} \tag{27}$$

Differentiating the above with respect to capital, and multiplying by the sector price, p_L, we obtain expressions for the value marginal product of capital in the L sector:

$$P_L MPK_L = P_L K_L^{-\beta} (A_L L)^{\beta} = P_L (1-\beta) \frac{Y_L}{K_L} \tag{28}$$

It is easy to show, using the expressions for P_H/P_L and K_H/K_L derived above, that this implies the value marginal product of capital is equal across sectors.

Further, note that when intermediate producers buy machines, they pay υR per unit of capital where υ is the markup over cost of producing machines (the rental rate). This implies that

$$\upsilon R = P_L MPK_L = P_L K_L^{-\beta} (A_L L)^{\beta} = P_L (1-\beta) \frac{Y_L}{K_L}$$

and we have

$$R = \frac{P_L MPK_L}{\upsilon} = P_L \left(\frac{1-\beta}{\upsilon}\right) \frac{Y_L}{K_L} = \left(\frac{1-\beta}{\upsilon}\right) \frac{Y}{K} \tag{29}$$

This last result, together with the fact that $r = (1-\tau)R - \delta$, where δ is the rate of depreciation of capital and τ is the tax on capital income

$$r = (1-\tau) \left(\frac{1-\beta}{\upsilon}\right) \frac{Y}{K} - \delta. \tag{30}$$

A.6. Innovation

Discovery of new blueprints for sector i is governed by the following process

$$\dot{A}_L = \eta_L \left(\frac{A_L^W}{A_L}\right)^{\varphi} \frac{Z_L}{N} \tag{31}$$

where represents A_L^W is the world frontier technology for sector L, η_L is the productivity of research effort, Z_L is the R&D expenditure on innovation or technology adoption in the L-sector, and φ measures the rate of technology diffusion (i.e. the strength of the benefit of the knowledge spillover from the world technology frontier).

In order to innovate, the entrepreneurs must incur an entry cost ζ, which is the same in both sectors and represents the costs of implementation/adaptation of new technology. Free entry into research implies that marginal benefit of extra innovation/adoption effort Z is equal to the cost or

$$\eta_L \left(\frac{A_L^W}{A_L}\right)^{\varphi} \frac{V_L}{N} = \zeta \tag{32}$$

where V_L is the value of a blueprint for a machine in sector L. Defining $\mu_L = \frac{A_L}{A_L^W}$ and dropping the country indicator, this equation implies that

$$\frac{V_H}{V_L} = \left(\frac{\eta_H}{\eta_L}\right)^{-1} \left(\frac{\mu_H}{\mu_L}\right)^{\varphi} \tag{33}$$

Finally, the value of a blueprint must satisfy the no-arbitrage condition

$$r\, V_L = \pi_L + \dot{V}_L \tag{34}$$

A.7. BGP growth rate & interest rate

Along the balanced growth path, the economy grows at a constant growth rate g, equal to the growth rate of the technology frontier (assumed to be the same for all types of skills).

$$g = \frac{1}{\theta}\left[r^* - \rho\right]$$

where ρ is the discount rate and θ is the CRRA coefficient. The BGP interest rate r^* therefore given by

$$r^* = \theta g + \rho, \tag{35}$$

and, using equation equations (29) and (30), the BGP rental rate is

$$R^* = \frac{\theta g + \rho + \delta}{1 - \tau} \tag{36}$$

Using the no-arbitrage conditions from (34) and the fact that along the BGP the value of a patent must be stationary ($\dot{V}_L = 0$), we get the following relationship between the value of a patent, profits and the interest rate

$$V_L = \frac{\pi_L}{r} \tag{37}$$

where profits are given by $\pi_i = \left(\frac{\upsilon-1}{\upsilon}\right) P_i^{1/\beta} N_i \left(\upsilon R^*\right)^{\frac{\beta-1}{\beta}}$. It follows that

$$\frac{V_H}{V_L} = \frac{\pi_H}{\pi_L} = \frac{\left(\frac{\upsilon-1}{\upsilon}\right) P_H^{*1/\beta} H \left(\upsilon R^*\right)^{\frac{\beta-1}{\beta}}}{\left(\frac{\upsilon-1}{\upsilon}\right) P_L^{*1/\beta} L \left(\upsilon R^*\right)^{\frac{\beta-1}{\beta}}} = \left(\frac{P_H}{P_L}\right)^{1/\beta} \frac{H}{L},$$

which can be further simplified using the expression for relative prices to obtain

$$\frac{V_H}{V_L} = \left(\frac{A_H}{A_L}\right)^{-\frac{1}{\sigma}} \left(\frac{H}{L}\right)^{\frac{\sigma-1}{\sigma}} \tag{38}$$

https://doi.org/10.1017/S1365100522000372 Published online by Cambridge University Press

Finally, combining equations (33) and (38) yields

$$\frac{A_H}{A_L} = \left(\frac{\eta_H}{\eta_L}\right)^{\frac{\sigma}{1+\varphi\sigma}} \left(\frac{H}{L}\right)^{\frac{\sigma-1}{1+\varphi\sigma}} \left(\frac{A_H^W}{A_L^W}\right)^{\frac{\varphi\sigma}{1+\varphi\sigma}} \tag{39}$$

Substituting the expression for relative productivity levels (39) into the relative wage formula (22), we obtain

$$\frac{w_H}{w_L} = \left(\frac{\eta_H}{\eta_L}\right)^{\frac{\sigma-1}{1+\varphi\sigma}} \left(\frac{H}{L}\right)^{\frac{\sigma-2-\varphi}{1+\varphi\sigma}} \left(\frac{A_H^W}{A_L^W}\right)^{\frac{\varphi(\sigma-1)}{1+\varphi\sigma}}. \tag{40}$$

From equations (32) and (37), we can see that on the BGP productivity relative to the frontier is given by

$$\mu_L = \left[\frac{\eta_L\left(\frac{\upsilon-1}{\upsilon}\right)(L/N)P_L^{*1/\beta}\left(\upsilon R^*\right)^{\frac{\beta-1}{\beta}}}{r^*\zeta}\right]^{1/\varphi}, \tag{41}$$

Appendix B: Transitional dynamics and the calibration of φ

We calibrate the value of φ to match the dynamic behavior of the model. Specifically, we choose a value of this parameter, which governs the strength of technology diffusion, to match the rate of convergence to the BGP. This section briefly describes the dynamics of our model.

Even with the assumption of constant supplies of skilled and unskilled labor (H and L), the dynamics of the model can be complicated. Because innovation for the two skill types and capital accumulation technologies are linear, the transitional dynamics may involve initial periods when only some of these activities take place. Eventually, the rates of return to all three activities are equalized and the economy converges to the BGP characterized in the paper. Characterizing the entire transitional dynamics of the model is beyond the scope of our analysis. Here we briefly discuss the dynamics of the system once all investment activities yield the same rate of return (and thus, all are undertaken). We show how to linearize the model around the BGP and discuss the implied speed of convergence which we use to choose the value for the diffusion parameter φ.

To characterize the dynamics of the model, we start by re-writing the free entry condition (where the equations are symmetric for the two skill types, we conserve space by presenting only one version)

$$V_H = \eta_H^{-1}\zeta H \mu_H^{\varphi} \tag{42}$$

we can differentiate the free entry condition to yield

$$\frac{\dot{V}_H}{V_H} = \varphi\,\frac{\dot{\mu}_H}{\mu_H} \tag{43}$$

Also since $\mu_H = A_H/A_H^W$, and the frontier is assumed to grow at the rate g, it follows from the expressions for the growth rate of productivity that

$$\frac{\dot{A}_H}{A_H} = \eta_H \mu_H^{-(1+\varphi)}\frac{\widetilde{Z}_H}{H} \tag{44}$$

and

$$\frac{\dot{A}_L}{A_L} = \eta_L \mu_L^{-\varphi}\frac{A_H^W \widetilde{Z}_L}{H} \tag{45}$$

https://doi.org/10.1017/S1365100522000372 Published online by Cambridge University Press

where $\widetilde{Z} \equiv Z/A_H^W$, which gives us the dynamic equations for the gaps to the frontier

$$\frac{\dot{\mu}_H}{\mu_H} = \eta_H \mu_H^{-(1+\varphi)} \frac{\widetilde{Z}_H}{H} - g \quad (46)$$

$$\frac{\dot{\mu}_L}{\mu_L} = \eta_L \mu_L^{-(1+\varphi)} \frac{A_H^W/A_L^W \widetilde{Z}_L}{H} - g \quad (47)$$

Additionally, recall that the no-arbitrage conditions are

$$\frac{\dot{V}_H}{V_H} = r - \frac{\pi_H}{V_H}$$

$$\frac{\dot{V}_L}{V_L} = r - \frac{\pi_L}{V_L}$$

Combining these conditions with the expression for profit rates derived earlier and equations (43), (46), and (47), we get

$$\widetilde{Z}_H = \eta_H^{-1} \mu_H^{1+\varphi} H \left(g + \left(r - \eta_H \left(\frac{\mu - 1}{\mu} \right) \frac{P_H^{1/\beta} (\mu R)^{\frac{\beta-1}{\beta}}}{\zeta \mu_H^{\varphi}} \right) / \varphi \right)$$

$$\widetilde{Z}_L = \frac{A_L^W}{A_H^W} \eta_L^{-1} \mu_L^{1+\varphi} H \left(g + \left(r - \eta_H \left(\frac{\mu - 1}{\mu} \right) \frac{P_L^{1/\beta} (H/L)^{-1} (\mu R)^{\frac{\beta-1}{\beta}}}{\zeta \mu_L^{\varphi}} \right) / \varphi \right)$$

Finally, using the budget constraint and the capital accumulation equation, we can derive the dynamics of K

$$I = Y - \zeta(Z_H + Z_L) - C$$

$$\frac{\dot{\widetilde{K}}}{\widetilde{K}} = I/K - \delta - g = \frac{\mu R}{1 - \beta} - \zeta \left(\frac{\widetilde{Z}_H}{\widetilde{K}} + \frac{\widetilde{Z}_L}{\widetilde{K}} \right) - \frac{\widetilde{C}}{\widetilde{K}} - \delta - g$$

The Euler equation completes the dynamical system

$$\frac{\dot{\widetilde{C}}}{\widetilde{C}} = \frac{1}{\theta}(r - \rho - \theta g)$$

The four difference equations in $\widetilde{C}, \widetilde{K}, \mu_H, \mu_L$ define the dynamics of the system. We linearize them around the BGP. It is tempting to view this system as one with one control ($\widetilde{C}$) and three state variables. However, recall that we have assumed that innovation and capital accumulation are all taking place (i.e. free entry conditions are binding), which, for a given value of initial physical capital, forces the values of μ_H, μ_L. This system only has one negative root and this root determines the speed of convergence to the BGP.

Setting all the other parameters equal to their calibrated values we let the speed of technology diffusion φ vary and calculate the speed of convergence to the BGP by solving the linearized system described above. Figure 3 plots the results. Our target value for the rate of convergence is 2.5% (Barro, 2012). Clearly, values of φ larger than 0.5 produce speeds of convergence well in excess of the target, while those much below it result in too slow convergence. We therefore choose our preferred value of φ to be 0.5.

Appendix C: The macro-approach vs. the conventional approach

To illustrate the difference between the macrointerpretation and the conventional one, consider Figure 4 below. It shows the error from following the standard way to uncover elasticity using

https://doi.org/10.1017/S1365100522000372 Published online by Cambridge University Press

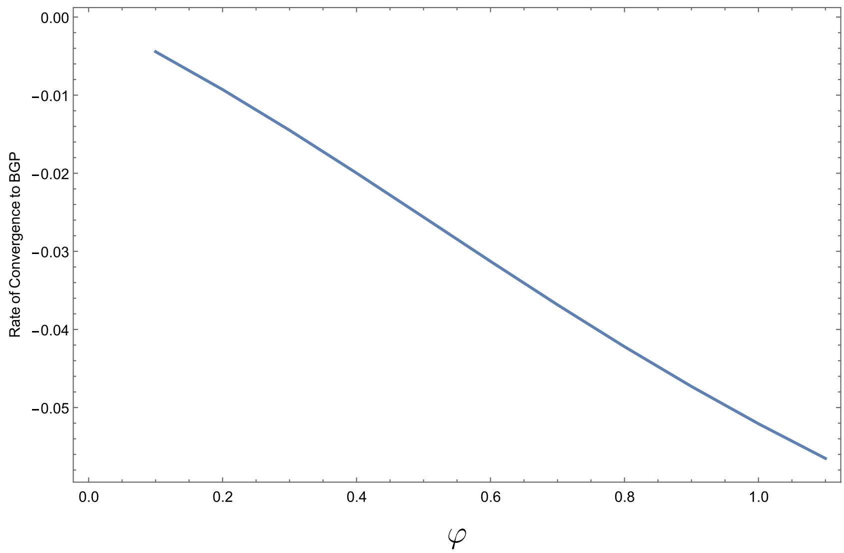


Figure 3. Speed of convergence to the BGP for different values of φ.

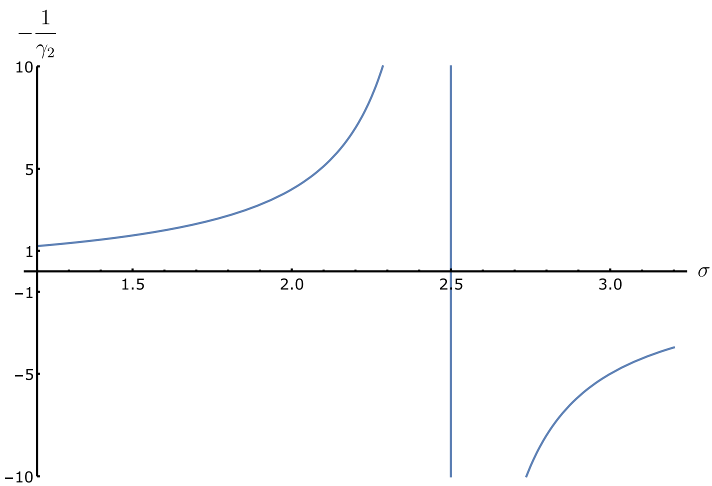


Figure 4. The error from following the standard way to uncover elasticity using estimates of equation (8). Vertical axis: the conventionally inferred elasticity (minus the inverse of the estimated coefficient on H/L in equation (8)). Horizontal axis: the true value of elasticity. Values plotted assume diffusion rate of $\varphi = 0.5$.

estimates of equation (8). Specifically, it plots the relationship between the true value of elasticity σ (on the horizontal axis) and the value that would be inferred if one consistently estimated equation (8) and used the negative value of the inverted coefficient on H/L (i.e. $-1/\gamma_2$) as the implied value of the elasticity, which is what the conventional approach does. (For this plot, the assumed value of the diffusion parameter for this plot is $\varphi = 0.5$, our preferred value). This procedure would clearly lead to an incorrect conclusion about the true value of elasticity and the error would be greater, the higher the value of elasticity. Importantly, if the true value of elasticity was close to the cutoff value for the strong skill bias (in the case of this plot, 2.5), the conventional procedure of uncovering elasticity would yield implausibly large positive or negative values, depending on

https://doi.org/10.1017/S1365100522000372 Published online by Cambridge University Press

whether the point estimate of γ_2 was slightly above or below its true value, either of which would, of course, occur in practice, even if the coefficient was estimated consistently.

Given the above discussion, we estimate elasticity by fitting equation (8) and inverting equation (9) to uncover the underlying value of σ. The standard errors associated with this estimate can be computed using the Delta method.

Cite this article: Jerzmanowski M and Tamura R (2023). "Aggregate elasticity of substitution between skills: estimates from a macroeconomic approach." *Macroeconomic Dynamics* **27**, 1597–1627. https://doi.org/10.1017/S1365100522000372