Aggregate Elasticity of Substitution between Skills: Estimates from a Macroeconomic Approach

Jerzmanowski, Michal and Tamura, Robert

Clemson University, Clemson University

28 May 2020

Aggregate Elasticity of Substitution between Skills: Estimates from a Macroeconomic Approach

Michał Jerzmanowski & Robert Tamura

May 30, 2020

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 U.S. micro data, 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 1.8 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 not implausible.

JEL classification: E24, E25, O11, J31

Keywords: elasticity of substitution, high-skill labor, low-skill labor, skill premium, strong skill-bias, endogenous directed technology

1 Introduction

Aggregate elasticity of substitution between workers with different skill levels is an important macroeconomics parameter. 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, Lochner, and Taber, 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 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 macro models 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.

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 micro evidence almost exclusively from the U.S. 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 macro panel 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 (for example, we have emerging economies like Poland alongside economies like Germany and the U.S. 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.

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 U.S.) 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 1.8 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, suggesting strong skill-bias is not implausible.

The paper is organized as follows. Section 1 discusses the canonical approach and the resulting estimating equation used in the literature based on the U.S. micro data. 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 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 micro data: 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.[1] 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 macro data 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

[1] Although, the addition of more flexible parametrization of the technology term brings that estimate down to about 1.8.

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.

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 (1996, 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 depend 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 conductivities 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 $i = H, L$. 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 & 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 time t (in practice, we use linear, quadratic, and country-specific trends) – becomes a regression equation

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

to be estimated using data on relative wages and relative labor supplies. 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. 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. The last fact 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 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 M. P. 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 U.S. 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 U.S. 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 (M. Timmer et al., 2007).[2] 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 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.[3] The potential for changes in definitions of skill groups over time leads the database authors to recommend the following

[2]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.

[3]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/`.

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.

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.

Table 1: Summary Statistics

Variable	Obs.	Mean	Std. Dev.	Min	Max
Share of High Skill Hours	609	0.23	0.09	0.05	0.48
Share of Middle Skill Hours	609	0.46	0.14	0.10	0.80
Share of Low Skill Hours	609	0.30	0.17	0.04	0.82
Share of High Skill Wages	609	0.33	0.11	0.06	0.63
Share of Middle Skill Wages	609	0.43	0.13	0.03	0.73
Share of Low Skill Wages	609	0.23	0.15	0.02	0.73
High Skill Wage/Lower Skill Wage	609	1.75	0.36	1.08	4.09
High Skill Wage/Lower Skill Hours	609	0.56	0.30	0.10	1.72
High Skill Wage/Middle Skill Wage	609	1.65	0.66	0.85	9.86
High Skill Wage/Middle Skill Hours	609	0.55	0.29	0.14	1.76

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-skill and lower-skill workers groups. We use OLS, fixed effects, and IV (where we instrument H/L with its lagged values) and for each estimation method 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 95% confidence 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 2.55, which is higher than most results in the literature but is in the plausible range and, in fact, not far of 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). 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 implausible value of -45. Note also, that the value of 2.55 implies the presence of *strong skill-bias* since equation (7) is satisfied.

When we use more flexible specifications and allow the time-trend to be country-specific, the point estimates of σ decrease to between 1.8 and 1.9 (columns 3-5). These are still higher than the conventionally accepted value of 1.6, but these estimates (and most of their 95% confidence intervals) fall outside the range needed for strong skill-bias. 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 considerably higher than 1.6, with values well above 2.0 within the confidence interval. And, when the country-specific trend squared is included, the regression implies no effect of changes in relative supply on relative wages, i.e. 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.

Table 2: High vs. Lower-skilled; OLS

$\log(H/L)$	0.022	0.023	-0.286	-0.289^{*}	-0.395^{***}
	(0.052)	(0.052)	(0.172)	(0.165)	(0.143)
Trend	*Yes*	*Yes*	*No*	*Yes*	*Yes*
Trend Squared	*No*	*Yes*	*No*	*Yes*	*Yes*
Country Trend	*No*	*No*	*Yes*	*Yes*	*Yes*
Country Trend Sq	*No*	*No*	*No*	*No*	*Yes*
σ	2.55	2.55	1.94	1.93	1.76
s.e.	(0.120)	(0.120)	(0.296)	(0.283)	(0.224)
95% Conf. Int.	[2.31, 2.79]	[2.31, 2.79]	[1.35, 2.53]	[1.37, 2.50]	[1.31, 2.21]
$p(\sigma < 2.5)$	0.34	0.33	0.95	0.96	1.00
σ'	-45.22	-43.56	3.50	3.46	2.53
R^2	0.01	0.01	0.66	0.67	0.89
N	609	609	609	609	609

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 $1/\gamma_2$.*

Table 3: High vs. Lower-skilled; Fixed Effects

	1	2	3	4	5
$\log(H/L)$	-0.285^{*}	-0.289^{*}	-0.393^{***}	-0.395^{***}	-0.062
	(0.166)	(0.159)	(0.130)	(0.137)	(0.113)
Trend	Yes	Yes	Yes	Yes	Yes
Trend Squared	No	Yes	No	Yes	Yes
Country Trend	No	No	Yes	Yes	Yes
Country Trend Sq	No	No	No	No	Yes
σ	1.94	1.93	1.76	1.76	2.36
s.e.	(0.286)	(0.274)	(0.204)	(0.215)	(0.240)
95% Conf. Int.	[1.37, 2.51]	[1.38, 2.48]	[1.35, 2.17]	[1.33, 2.19]	[1.88, 2.84]
p($\sigma < 2.5$)	0.95	0.96	1.00	1.00	0.71
σ'	3.50	3.46	2.55	2.53	16.05
R^2	0.18	0.21	0.72	0.73	0.86
N	609	609	609	609	609

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* $1/\gamma_2$*.*

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, ignoring technology diffusion cross countries, would lead to estimates of elasticity that are implausible and very sensitive to the time specification.

Table 4: High vs. Lower-skilled; IV

	1	2	3	4	5
$\log(H/L)$	-0.307^{**}	-0.318^{**}	-0.409^{***}	-0.416^{**}	-0.146
	(0.151)	(0.138)	(0.150)	(0.169)	(0.131)
Trend	Yes	Yes	No	Yes	Yes
Trend Squared	No	Yes	Yes	Yes	Yes
Country Trend	No	No	Yes	Yes	Yes
Country Trend Sq	No	No	No	No	Yes
σ	1.90	1.88	1.74	1.73	2.19
s.e.	(0.256)	(0.232)	(0.232)	(0.260)	(0.256)
95% Conf. Int.	[1.39, 2.41]	[1.42, 2.35]	[1.27, 2.20]	[1.20, 2.25]	[1.68, 2.71]
$p(\sigma < 2.5)$	.	.	.	.	.
σ'	3.26	3.14	2.45	2.40	6.85
R^2	0.21	0.27	0.74	0.74	0.87
N	501	501	501	501	501

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. 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* $1/\gamma_2$*.*

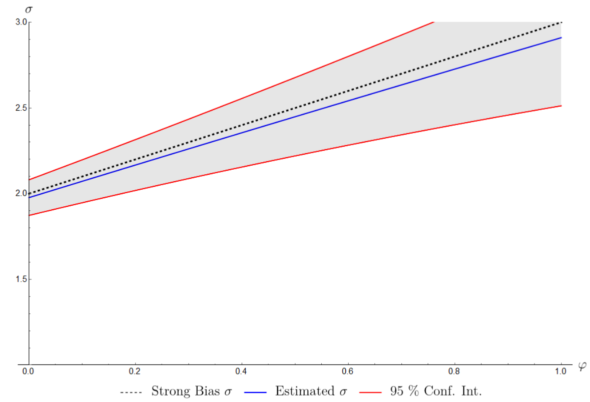
We conclude that our estimates, together with our macro interpretation, which accounts for directed technology change and cross-country idea diffusion, produce a set of estimates of the elasticity of substitution that are higher than those found using U.S. micro data but fall within a plausible range. 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 macro interpretation 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.

4.2 Robustness

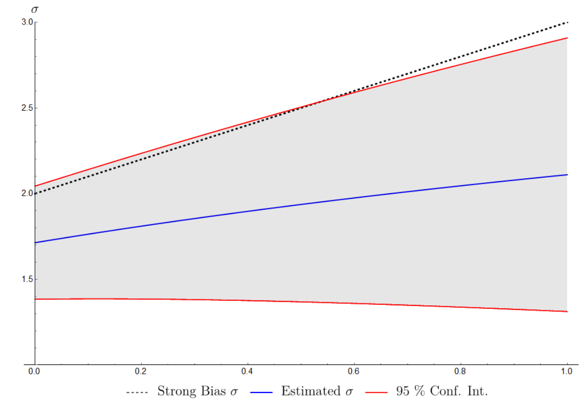
Different Rates of Diffusion (φ)

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 95% confidence 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 95% confidence band, and the black line is the value of elasticity above which the strong skill bias is present.

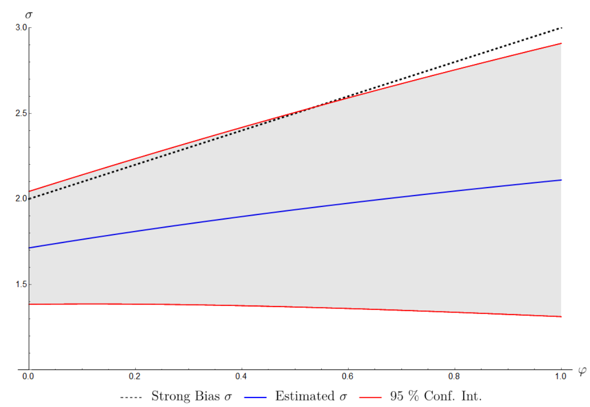
As can be seen from the graphs, given the estimated regression coefficients, the implied elasticity is increasing in φ, but, of course, so is the cut-off for strong skill-bias. In the case of OLS, point estimates are always only slightly below the threshold when a country-specific linear trend is not included (panel 1(a)) but fall farther from the threshold once this variable is included in the regression. Regardless of the specification, the elasticity consistent with strong bias falls within the 95% confidence interval either always (without country-specific trend) or for $\varphi < 0.6$. Things look similar under fixed effects, except the point estimates are lower and, as a result, when a country-specific trend is included, strong skill bias is within the 95% confidence band only for diffusion rates lower than our preferred value of 0.5. 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.



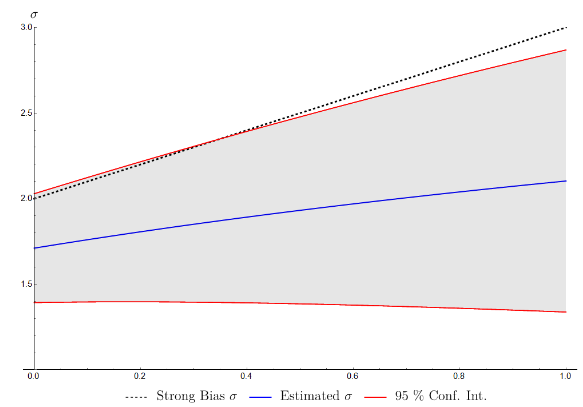
(a) OLS: specification with common linear and quadratic trends (column 2 in table 1)



(b) OLS: specification with common linear and quadratic trends and country-specific linear trend (column 4 in table 1)



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



(d) FE: specification with common linear and quadratic trends and country-specific linear trend (column 4 in table 2)

Figure 1: Estimates of σ (blue line) and the 95% confidence 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.

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 M. P. 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-skill group as the lower-skilled workers. The results are in Tables 5 - 12. The point estimates are mostly higher but still within the plausible range. When country fixed effects are included, estimated elasticity is considerably higher than 1.6, with six out of ten specifications yielding point estimates above two and 95% values above the strong skill-bias threshold (Tables 6 -12). Again, when country-level effects and country-specific linear and quadratic trends are included, the point estimate is itself consistent with the presence of strong bias. The conventionally computed elasticities are again considerably higher, and in most cases implausibly so (Tables 6 - 12).

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 8 - 13. They are generally quite similar to our baseline results

Table 5: High vs. Middle-skilled; OLS

	1	2	3	4	5
$\log(H/L)$	0.014	0.018	-0.200	-0.146	-0.442^{**}
	(0.101)	(0.100)	(0.119)	(0.116)	(0.214)
Trend	Yes	Yes	Yes	Yes	Yes
Trend Squared	No	Yes	No	Yes	Yes
Country Trend	No	No	Yes	Yes	Yes
Country Trend Sq	No	No	No	No	Yes
σ	2.53	2.54	2.09	2.19	1.68
s.e.	(0.231)	(0.230)	(0.222)	(0.227)	(0.322)
95% Conf. Int.	[2.07, 2.99]	[2.08, 3.00]	[1.65, 2.54]	[1.74, 2.65]	[1.04, 2.33]
p($\sigma < 2.5$)	0.45	0.43	0.95	0.89	0.98
σ'	-71.45	-54.27	5.01	6.83	2.26
R^2	0.00	0.01	0.56	0.57	0.90
N	609	609	609	609	609

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*. 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 6: High vs. Middle-skilled; Fixed Effects

	1	2	3	4	5
$\log(H/L)$	-0.208^{*}	-0.155	-0.346^{***}	-0.441^{**}	0.052
	(0.121)	(0.118)	(0.111)	(0.206)	(0.162)
Trend	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Trend Squared	*No*	*Yes*	*No*	*Yes*	*Yes*
Country Trend	*No*	*No*	*Yes*	*Yes*	*Yes*
Country Trend Sq	*No*	*No*	*No*	*No*	*Yes*
σ	2.08	2.18	1.84	1.69	2.62
s.e.	(0.223)	(0.228)	(0.182)	(0.311)	(0.384)
95% Conf. Int.	[1.63, 2.52]	[1.72, 2.63]	[1.47, 2.20]	[1.07, 2.31]	[1.85, 3.39]
p(σ < 2.5)	0.95	0.90	1.00	0.98	0.37
σ'	4.81	6.47	2.89	2.27	-19.17
R^2	0.07	0.10	0.79	0.79	0.92
N	609	609	609	609	609

NOTES: Estimates of $\log\left(w_H/w_L\right)_{it} = \lambda_i + \gamma_0 + \gamma_1\, \mu(t) + \gamma_2 \log\left(H/L\right)_{it} + \varepsilon_{it}$, *using high-skilled workers as* H *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* $1/\gamma_2$.

Table 7: High vs. Middle-skilled; IV

	1	2	3	4	5
$\log(H/L)$	-0.196^{*}	-0.133	-0.354^{***}	-0.409^{**}	0.044
	(0.109)	(0.106)	(0.115)	(0.202)	(0.257)
Trend	Yes	Yes	Yes	Yes	Yes
Trend Squared	No	Yes	No	Yes	Yes
Country Trend	No	No	Yes	Yes	Yes
Country Trend Sq	No	No	No	No	Yes
σ	2.10	2.22	1.82	1.74	2.60
s.e.	(0.204)	(0.210)	(0.187)	(0.313)	(0.605)
95% Conf. Int.	[1.69, 2.51]	[1.80, 2.64]	[1.45, 2.20]	[1.11, 2.36]	[1.39, 3.81]
p(σ < 2.5)	.	.	.	.	.
σ'	5.10	7.50	2.82	2.44	-22.50
R^2e	0.08	0.12	0.80	0.80	0.91
N	501	501	501	501	501

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 8: High vs. Lower-skilled ; OLS (1970-2005)

	1	2	3	4	5
$\log(H/L)$	0.001	0.001	-0.357	-0.357	-0.249^{**}
	(0.065)	(0.064)	(0.205)	(0.205)	(0.097)
Trend	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Trend Squared	*No*	*Yes*	*No*	*Yes*	*Yes*
Country Trend	*No*	*No*	*Yes*	*Yes*	*Yes*
Country Trend Sq	*No*	*No*	*No*	*No*	*Yes*
σ	2.50	2.50	1.82	1.82	2.00
s.e.	(0.146)	(0.144)	(0.331)	(0.332)	(0.173)
95% Conf. Int.	[2.21, 2.79]	[2.21, 2.79]	[1.16, 2.48]	[1.15, 2.48]	[1.66, 2.35]
$p(\sigma < 2.5)$	0.49	0.50	0.95	0.95	0.99
σ'	-1092.98	-1352.06	2.80	2.80	4.02
R^2	0.01	0.01	0.72	0.72	0.95
N	400	400	400	400	400

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 9: High vs. Lower-skilled ; Fixed Effects (1970-2005)

	1	2	3	4	5
$\log(H/L)$	-0.359^{*}	-0.359^{*}	-0.248^{**}	-0.251^{**}	-0.262^{*}
	(0.199)	(0.199)	(0.096)	(0.094)	(0.123)
Trend	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Trend Squared	*No*	*Yes*	*No*	*Yes*	*Yes*
Country Trend	*No*	*No*	*Yes*	*Yes*	*Yes*
Country Trend Sq	*No*	*No*	*No*	*No*	*Yes*
σ	1.81	1.82	2.00	2.00	1.98
s.e.	(0.322)	(0.323)	(0.171)	(0.168)	(0.216)
95% Conf. Int.	[1.17, 2.46]	[1.17, 2.46]	[1.66, 2.34]	[1.66, 2.33]	[1.55, 2.41]
$p(\sigma < 2.5)$	0.95	0.95	0.99	0.99	0.97
σ'	2.78	2.78	4.03	3.99	3.82
R^2	0.23	0.23	0.86	0.87	0.94
N	400	400	400	400	400

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 10: High vs. Lower-skilled ; IV (1970-2005)

	1	2	3	4	5
log(H/L)	-0.355^{*}	-0.355^{*}	-0.227^{**}	-0.227^{**}	-0.265^{**}
	(0.201)	(0.201)	(0.094)	(0.094)	(0.129)
Trend	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Trend Squared	*No*	*Yes*	*No*	*Yes*	*Yes*
Country Trend	*No*	*No*	*Yes*	*Yes*	*Yes*
Country Trend Sq	*No*	*No*	*No*	*No*	*Yes*
σ	1.82	1.82	2.04	2.04	1.97
s.e.	(0.326)	(0.326)	(0.170)	(0.170)	(0.226)
95% Conf. Int.	[1.17, 2.47]	[1.17, 2.47]	[1.70, 2.38]	[1.70, 2.38]	[1.52, 2.43]
p($\sigma < 2.5$)	0.97	0.97	0.98	0.98	0.96
σ'	2.82	2.82	4.41	4.41	3.77
R^2	0.25	0.25	0.87	0.87	0.95
N	384	384	384	384	384

NOTES: IV estimates of $\log\left(w_H/w_L\right)_{it} = \lambda_i + \gamma_0 + \gamma_1\,\mu(t) + \gamma_2 \log\left(H/L\right)_{it} + \varepsilon_{it}$*, 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* $1/\gamma_2$*.*

5 Conclusions

Aggregate elasticity of substitution between workers with different skill levels is an important macroeconomics parameter. 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 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 U.S. micro data, lead 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 it is not clear whether using an elasticity estimate obtained using U.S. micro data is suitable for calibrating models aimed at explaining the behavior of widely diverse groups of economies (Jones, 2014) or if assumptions about constant human capital quality across a long time span – implied when using long U.S. micro time-series – are justified (Bowlus et al., 2017). We contribute to this literature by estimating the elasticity of substitution between workers of different skill types using a macro panel 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 U.S. 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 1.8 and 2.6, 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. Finally, 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, suggesting strong skill-bias is not implausible.

References

Acemoglu, Daron (1998). "Why do new technologies complement skills? Directed Technical change and wage inequality". In: *Quarterly Journal of Economics* 113, pp. 1055–1090.

— (2002). "Directed technical change". In: *Review of Economic Studies* 69, pp. 781–809.

Acemoglu, Daron and David Autor (2011). "Skills, Tasks and Technologies: Implications for Employment and Earnings". In: *Handbook of Labor Economics*. Ed. by P O. Ashenfelter and D. Card. Vol. 4. Elsevier B.V. Chap. 12, pp. 1043–1171.

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].

Barro, Robert J (Aug. 2012). *Convergence and Modernization Revisited*. Working Paper 18295. National Bureau of Economic Research. DOI: 10.3386/w18295. URL: http://www.nber.org/papers/w18295.

Bowlus, Audra, Eda Bozkurt, Lance Lochner, and Chris Robinson (Nov. 2017). *Wages and Employment: The Canonical Model Revisited*. Working Paper 24069. National Bureau of Economic Research. DOI: 10.3386/w24069. URL: http://www.nber.org/papers/w24069.

Caselli, Francesco and Antonio Ciccone (Mar. 2019). "The Human Capital Stock: A Generalized Approach: Comment". In: *American Economic Review* 109.3, pp. 1155–1174. URL: https://ideas.repec.org/a/aea/aecrev/v109y2019i3p1155-74.html.

Caselli, Francesco and Wilbur John Coleman (2006). "The World Technology Frontier". In: *American Economic Review* 96, pp. 499–522.

Ciccone, Antonio and Giovanni Peri (2005). "The Long-Run Substitutability Between More and Less Educated Workers: Evidence from U.S. States, 1950-1990". In: *The Review of Economics and Statistics* 87, pp. 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, William H. (2003). *Econometric Analysis*. Fifth. Pearson Education. ISBN: 0-13-066189-9. URL: http://pages.stern.nyu.edu/~wgreene/Text/econometricanalysis.htm.

Heckman, James, Lance Lochner, and Christopher Taber (Jan. 1998). "Explaining Rising Wage Inequality: Explanations With A Dynamic General Equilibrium Model of Labor Earnings With Heterogeneous Agents". In: *Review of Economic Dynamics* 1.1, pp. 1–58. DOI: 10.1006/redy.1997.0008. URL: https://ideas.repec.org/a/red/issued/v1y1998i1p1-58.html.

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

Jerzmanowski, Michal and Robert Tamura (2019). "Directed technological change & cross-country income differences: A quantitative analysis". In: *Journal of Development Economics* 141.C. DOI: 10.1016/j.jdeveco.2019.10. URL: https://ideas.repec.org/a/eee/deveco/v141y2019ics0304387818305820.html.

Johnson, G. (1970). "The Demand For Labor by Educational Category". In: *Southern Economic Journal* 37, pp. 190–204.

Jones, Benjamin F. (Nov. 2014). "The Human Capital Stock: A Generalized Approach". In: *American Economic Review* 104.11, pp. 3752–77. DOI: 10.1257/aer.104.11.3752. URL: http://www.aeaweb.org/articles?id=10.1257/aer.104.11.3752.

Katz, Lawrence F. and Kevin M. Murphy (1992). "Changes in Relative Wages, 1963–1987: Supply and Demand Factors". In: *The Quarterly Journal of Economics* 107.1, pp. 35–78. URL: https://ideas.repec.org/a/oup/qjecon/v107y1992i1p35-78..html.

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

Krusell, Per, Lee E. Ohanian, Jose-Victor Rios-Rull, and Giovanni L. Violante (Sept. 2000). "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis". In: *Econometrica* 68(5), pp. 1029–1054.

O'Mahony, Mary and Marcel P. Timmer (June 2009). "Output, Input And Productivity Measures At The Industry Level: The EU KLEMS Database". In: *The Economic Journal* 119, pp. 374–403.

Tamura, Robert, Jerry Dwyer, John Devereux, and Scott Baier (2019). "Economic growth in the long run". In: *Journal of Development Economics* 137.C, pp. 1–35. DOI: 10.1016/j.jdeveco.2018.10. URL: https://ideas.repec.org/a/eee/deveco/v137y2019icp1-35.html.

Timmer, Marcel, Ton van Moergastel, Edwin Stuivenwold, Gerard Ypma, Mary O'Mahony, and Mari 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].

Appendices

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.

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.

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 numereire so that

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

Intermediate Goods & Machines

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

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

$$\tag{13}$$

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:

$$\max_{\{\chi_{jL},L\}} \{\frac{P_L}{1-\beta} \int_0^{A_L} \chi_{jL}^{1-\beta} dj L^\beta - \int_0^{A_L} p_{jL}\chi_{jL} dj - w_L L\}, \tag{14}$$

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

$$P_L \chi_{jL}^{-\beta} L^\beta = p_{jL} \tag{15}$$

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. 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.

$$p_{jL} = \upsilon R \tag{16}$$

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

$$\chi_{jL} = \left(\frac{P_L}{\upsilon R}\right)^{1/\beta} L \tag{17}$$

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

$$Y_L = \frac{1}{1-\beta}\left(\frac{P_L}{\upsilon R}\right)^{\frac{1-\beta}{\beta}} A_L L \tag{18}$$

And the profit per line of machines is given by

$$\pi_{jL} = \left(\frac{\upsilon - 1}{\upsilon}\right) P_L^{1/\beta} L \left(\upsilon R\right)^{\frac{\beta-1}{\beta}} \tag{19}$$

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

$$\frac{P_H}{P_L} = \left(\frac{A_H H}{A_L L}\right)^{-\frac{\beta}{\sigma}} \tag{20}$$

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

Wages & Technology

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

$$\frac{\beta P_L}{1-\beta}\int_0^{A_L} \chi_{jL}^{1-\beta} dj L^{\beta-1} = w_L, \tag{21}$$

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

$$w_L = \frac{\beta}{1-\beta} A_L \beta P_L^{\frac{1}{\beta}} (\upsilon R)^{-\frac{1-\beta}{\beta}} \tag{22}$$

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

$$\frac{w_H}{w_L} = \left(\frac{A_H}{A_L}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{H}{L}\right)^{-\frac{1}{\sigma}} \tag{23}$$

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

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

$$K_L = \int_0^{A_L} \chi_{jL} dj = A_L \left(\frac{P_L}{\upsilon R}\right)^{1/\beta} L \tag{24}$$

and it follows that

$$\frac{K_H}{K_L} = \left(\frac{A_H}{A_L}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{H}{L}\right)^{\frac{\sigma-1}{\sigma}} \tag{25}$$

In addition, total capital stock by K is given by

$$K = K_L + K_H. \tag{26}$$

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

$$K_L = A_L \chi_L \tag{27}$$

$$\tag{28}$$

Substituting this into the intermediate goods production functions

$$Y_L = \frac{1}{1-\beta} K_L^{1-\beta} (A_L L)^{\beta} \tag{29}$$

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:

$$P_L MPK_L = P_L K_L^{-\beta} (A_L L)^{\beta} = P_L (1-\beta) \frac{Y_L}{K_L} \tag{30}$$

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 mark-up 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

$$R = \frac{P_L MPK_L}{\upsilon} = P_L \left(\frac{1-\beta}{\upsilon}\right) \frac{Y_L}{K_L} = \left(\frac{1-\beta}{\upsilon}\right) \frac{Y}{K} \tag{31}$$

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

$$r = (1-\tau)\left(\frac{1-\beta}{\upsilon}\right) \frac{Y}{K} - \delta. \tag{32}$$

Innovation

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

$$\dot{A}_L = \eta_L \left(\frac{A_L^W}{A_L}\right)^{\varphi} \frac{Z_L}{N} \tag{33}$$

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). [4]

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

$$\eta_L \left(\frac{A_L^W}{A_L}\right)^{\varphi} \frac{V_L}{N} = \zeta \tag{34}$$

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

$$\frac{V_H}{V_L} = \left(\frac{\eta_H}{\eta_L}\right)^{-1} \left(\frac{\mu_H}{\mu_L}\right)^{\varphi} \tag{35}$$

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

[4] In Tamura et al. (2019) the authors show that an intergenerational human capital model with conditional diffusion of human capital can explain both the average growth rate of output per worker as well as the cross sectional variation in the average growth rate of output per worker. Here we abstract from the dependence of the rate of diffusion on human capital, although we view this as a potentially interesting extension."

$$r\,V_L = \pi_L + \dot{V}_L \tag{36}$$

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

$$r^* = \theta g + \rho, \tag{37}$$

and, using equation equations (31) and (32), the BGP rental rate is

$$R^* = \frac{\theta g + \rho + \delta}{1-\tau} \tag{38}$$

Using the no-arbitrage conditions from (36) 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

$$V_L = \frac{\pi_L}{r} \tag{39}$$

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

$$\frac{V_H}{V_L} = \left(\frac{A_H}{A_L}\right)^{-\frac{1}{\sigma}} \left(\frac{H}{L}\right)^{\frac{\sigma-1}{\sigma}} \tag{40}$$

Finally, combining equations (35) and (40), yields

$$\frac{A_H}{A_L} = \left(\frac{\eta_H}{\eta_L}\right)^{\frac{\sigma}{1+\varphi\sigma}} \left(\frac{H}{L}\right)^{\frac{\sigma-1}{1+\varphi\sigma}} \left(\frac{A_H^W}{A_L^W}\right)^{\frac{\varphi\sigma}{1+\varphi\sigma}} \tag{41}$$

Substituting the expression for relative productivity levels (41) into the relative wage formula (23) we obtain

$$\frac{w_H}{w_L} = \left(\frac{\eta_H}{\eta_L}\right)^{\frac{\sigma-1}{1+\varphi\sigma}} \left(\frac{H}{L}\right)^{\frac{\sigma-2-\varphi}{1+\varphi\sigma}} \left(\frac{A_H^W}{A_L^W}\right)^{\frac{\varphi(\sigma-1)}{1+\varphi\sigma}} . \tag{42}$$

From equations (34) and (39) we can see that on the BGP productivity relative to the frontier is given by

$$\mu_L = \left[\frac{\eta_L \left(\frac{\upsilon-1}{\upsilon}\right)(L/N) P_L^{*1/\beta} \left(\upsilon R^*\right)^{\frac{\beta-1}{\beta}}}{r^*\zeta}\right]^{1/\varphi} , \tag{43}$$

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)

$$V_H = \eta_H^{-1} \zeta H \mu_H^{\varphi} \tag{44}$$

we can differentiate the free entry condition to yield

$$\frac{\dot{V}_H}{V_H} = \varphi \frac{\dot{\mu}_H}{\mu_H} \tag{45}$$

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

$$\frac{\dot{A}_H}{A_H} = \eta_H \mu_H^{-(1+\varphi)} \frac{\widetilde{Z}_H}{H} \tag{46}$$

and

$$\frac{\dot{A}_L}{A_L} = \eta_L \mu_L^{-\varphi} \frac{A_H^W \widetilde{Z}_L}{H} \tag{47}$$

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

$$\frac{\dot{\mu}_H}{\mu_H} = \eta_H \mu_H^{-(1+\varphi)} \frac{\widetilde{Z}_H}{H} - g \tag{48}$$

$$\frac{\dot{\mu}_L}{\mu_L} = \eta_L \mu_L^{-(1+\varphi)} \frac{A_H^W / A_L^W \widetilde{Z}_L}{H} - g \tag{49}$$

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 (45), (48), and (49) 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 2 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.

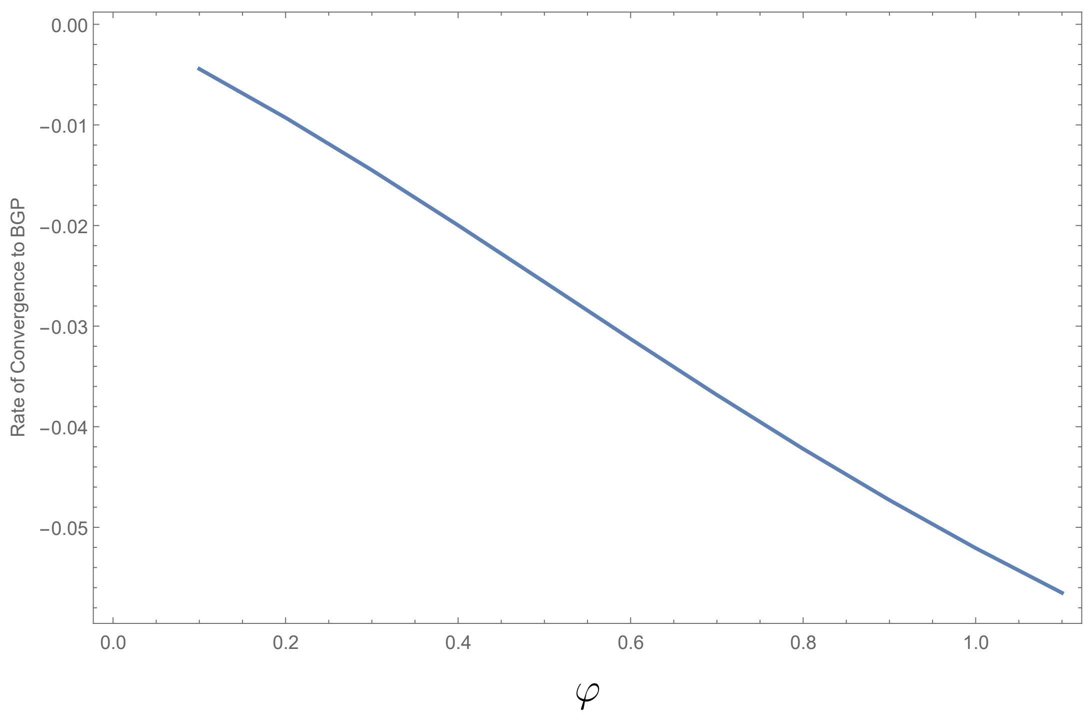


Figure 2: Speed of convergence to the BGP for different values of φ.

C GMM Results

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). We do not have a good candidate for an instrument, but in Table 4 of the paper we report the estimate of our model using lagged values of relative labor supplies as instruments for the current level. Here, in Tables 11-13 we report estimating the same model using an Arellano-Bond type GMM estimator. These results are very similar to those from the fixed effects and IV models above.

Table 11: High vs. Lower-skilled ; GMM

	1	2	3	4	5
$\log(H/L)$	-0.291	-0.290^{*}	-0.399^{***}	-0.404^{**}	-0.100
	(0.175)	(0.162)	(0.139)	(0.148)	(0.108)
Trend	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Trend Squared	*No*	*Yes*	*No*	*Yes*	*Yes*
Country Trend	*No*	*No*	*Yes*	*Yes*	*Yes*
Country Trend Sq	*No*	*No*	*No*	*No*	*Yes*
σ	1.93	1.93	1.75	1.74	2.28
s.e.	(0.300)	(0.278)	(0.218)	(0.230)	(0.220)
95% Conf. Int.	[1.33, 2.53]	[1.37, 2.49]	[1.32, 2.19]	[1.28, 2.20]	[1.85, 2.72]
$p(\sigma < 2.5)$	0.95	0.96	1.00	0.99	0.82
σ'	3.44	3.44	2.51	2.47	9.96
N	573	573	573	573	573

NOTES: 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 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 12: High vs. Middle-skilled; GMM

	1	2	3	4	5
$\log(H/L)$	-0.143	-0.050	-0.359^{***}	-0.444^{**}	0.016
	(0.095)	(0.065)	(0.118)	(0.206)	(0.165)
Trend	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Trend Squared	*No*	*Yes*	*No*	*Yes*	*Yes*
Country Trend	*No*	*No*	*Yes*	*Yes*	*Yes*
Country Trend Sq	*No*	*No*	*No*	*No*	*Yes*
σ	2.20	2.39	1.81	1.68	2.54
s.e.	(0.185)	(0.140)	(0.191)	(0.311)	(0.377)
95% Conf. Int.	[1.83, 2.57]	[2.11, 2.67]	[1.43, 2.20]	[1.06, 2.30]	[1.78, 3.29]
p($\sigma < 2.5$)	0.93	0.78	1.00	0.98	0.46
σ'	7.00	19.88	2.78	2.25	-63.50
N	573	573	573	573	573

NOTES: 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 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. Lower-skilled ; GMM (1970-2005)

	1	2	3	4	5
$\log(H/L)$	-0.355^{*}	-0.358^{*}	-0.234^{**}	-0.231^{**}	-0.267^{*}
	(0.189)	(0.187)	(0.103)	(0.101)	(0.128)
Trend	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Trend Squared	*No*	*Yes*	*No*	*Yes*	*Yes*
Country Trend	*No*	*No*	*Yes*	*Yes*	*Yes*
Country Trend Sq	*No*	*No*	*No*	*No*	*Yes*
σ	1.82	1.82	2.03	2.03	1.97
s.e.	(0.306)	(0.303)	(0.186)	(0.182)	(0.225)
95% Conf. Int.	[1.21, 2.43]	[1.21, 2.42]	[1.66, 2.40]	[1.67, 2.40]	[1.52, 2.42]
p($\sigma < 2.5$)	0.96	0.96	0.98	0.98	0.97
σ'	2.82	2.79	4.26	4.33	3.75
N	384	384	384	384	384

NOTES: 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 *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* $1/\gamma_2$*.*