Discussion Papers in Economics

UNRAVELING THE SKILL PREMIUM

By

Peter McAdam
(European Central Bank and University of Surrey)
&
AlpoWillman
(University of Kent)

DP 01/17

School of Economics
University of Surrey
Guildford
Surrey GU2 7XH, UK
Telephone +44 (0)1483 689380
Facsimile +44 (0)1483 689548
Web www.econ.surrey.ac.uk
ISSN: 1749-5075

UNRAVELING THE SKILL PREMIUM

Peter McAdam
European Central Bank and University of Surrey
peter.mcadam@ecb.europa.eu

Alpo Willman
University of Kent

We thank two anonymous referees for helpful comments as well as Robert Chirinko, Fabrizio Colonna (discussant) and participants at numerous seminars. The views expressed are not necessarily those of the ECB. Address correspondence to: Peter McAdam, Research Department, European Central Bank, Sonnemannstraße 60314, Frankfurt am Main, Germany; e-mail: peter.mcadam@ecb.europa.eu.

Abstract

For the US the supply and wages of skilled labor relative to those of unskilled labor have grown over the postwar period. The literature has tended to explain this through "skill-biased technical change". Empirical work has concentrated around two variants: (1) Capital-skill complementarity, (2) Skill-augmenting technical change. Our purpose is to nest and discriminate between these two explanations. We do so in the framework of multi-level CES production function where factors are disaggregated into skilled and unskilled labor, and capital into structures and equipment capital. Using a 5-equation system approach and several nesting alternatives, we retrieve estimates of the substitution elasticities and technical changes. Our estimations can produce results in line with capital-skill-complementarity hypothesis. However, those results are outperformed where the only source of the widening skill-premium has been skill-augmenting technical change. We also show that the different explanations for SBTC have different implications for projected developments of the premium.

Keywords: Skill Premium, Inequality, Multi-level CES production function,
Factor-Augmenting Technical Progress, Capital-Skill Complementarity.

1 INTRODUCTION

A widely documented fact for the US is that both the supply *and* wages of skilled labor relative to those of unskilled labor have grown markedly over the postwar period, see Bound and Johnson (1992), Goldin and Katz (2010), Acemoglu and Autor (2011). The literature has tended to explain this growing wage inequality by *skill-biased technical change* (SBTC). This is defined as a change in the production technology that favors skilled over unskilled labor by increasing its relative productivity and, hence, relative demand. Empirical work has concentrated around two variants of SBTC:[1]

(i) Capital-skill complementarity (KSC), and;

(ii) Technical change (TC) that augments skilled more than unskilled labor.

In our exercises we nest these two (not necessarily mutually exclusive) explanations in a multi-equation production and technology system. In doing so, our purpose is to understand the differences between them and econometrically discriminate between them. On the first explanation, Griliches (1969), showed that, for US manufacturing, capital and skilled labor were more complementary than capital and unskilled labor. The hypothesis gained renewed interest given the decline in the constant-quality relative price of equipment, e.g. Gordon (1990). This decline expanded the use of such capital, Autor et al. (1998), which, given the complementarity structure, increased the relative demand for skilled labor *and* – despite the latter's increased supply – a persistent rise in the skill premium (Greenwood et al. (1997), Krusell et al. (2000)).

The second hypothesis, following Tinbergen (1974) (discussed and extended by Katz and Murphy (1992), Acemoglu (2002), Autor et al. (2008)), is concentrated in the substitution between skilled and unskilled labor *and* the growth rates of "factor-augmenting" technical changes. Under this hypothesis an essential starting point is that, although skilled and unskilled labor are imperfect substitutes, through a Constant Elasticity of Substitution (CES) aggregator they form an aggregate quantity index separable from capital inputs. Now an above-unity substitution elasticity between skilled and unskilled labor, coupled with faster skilled than unskilled labor-augmenting technical change, results in the skill biased technical change and the widening skill premium. However, the assumed underlying separability is not tested and therefore those results do not suffice to validate skill-augmenting technical change as the driver of skill-biased technical change.

Both approaches, note, aim at explaining the same observed wage premium. Both, as we shall demonstrate, rely on particular production nestings and configurations for the substitution elasticities and factor-augmenting technical progress rates. Notwithstanding these conceptual similarities, though, they are fundamentally different explanations. KSC refers to the curvature of the isoquants and thus the ease with which factors can be substituted for one another. TC captures non-parallel shifts of the isoquants.

Unravelling and understanding their separate effects is our objective. And, given their apparent observational equivalence, this should be done in a robust and rigorous manner. In doing so, we provide the following main contributions to the literature.

First, we generalize the Krusell et al. (2000) four-factor production framework into, nested three- and two-level CES functions with factor-augmenting technology. We derive the five non-linear equation system with cross-equation constraints (i.e., four FOCs plus production function). We estimate them under different assumptions on how the four factors, and technical change augmenting them, are combined in three- and two-level CES production functions. As will become clear, the key advantage of this approach is that it offers a unified analysis, containing *both* alternative hypothesis of biased technical change as special cases. Moreover, this system approach models not only the skill premium but also the individual wages of skilled and unskilled labor, potential output, and the user costs of the two capital types. This constitutes a more rigorous and falsifiable representation of the data that has so far been attempted.

Our second contribution is that our framework allows us to reduce the development of the skill premium to originate from three separate sources: KSC, TC and the "relative supply of skills" (RSS). Arguably, the literature (and practitioners) tend to see the first two as distinct phenomena. Our decomposition, however, demonstrates that all channels may simultaneously be working in determining the evolution of the premium; over time, different channels may complement or offset one another.

These three terms have slightly different meanings depending on the production-technology system used. But broadly speaking we can classify them as follows.

(1) The RSS relates to how the growth of unskilled workers and skilled workers impact the premia. In our data sample, the growth of the latter exceeds the former. In itself, one might conclude that this would reduce the skill premia. However, whether the RSS channel impacts the premia positively or negatively depends on the associated elasticity of between and across skilled versus unskilled labor and different capital types, as well as the values of the distribution parameters plus (as we explain) the form of the production system.

(2) Second, there is the capital skill complementarity (KSC) effect itself. This, analogously, relates to the growth of both types of capital. The growth of both can impact positively or negatively the skill premium depending on the degrees to which either capital type are complementary to skilled labor. (3) Finally, there is technical change (TC). This term gathers together all the factor-augmenting technology terms weighted

by the parameters of the system. Depending on these (often complicated sets of parameter combinations), this effect can be positive or negative. For the various estimated cases, we can show their implications for the future development of the premium and other variables of interest under alternative assumptions of technical change and the projected growth rates of input factors. Again, what different explanations for SBTC imply for future earnings and growth appears to have been neglected in the academic literature despite its clear policy relevance, e.g., Dobbs et al. (2012).

Our main conclusion is that capital skill complementarity can be found in the data under appropriately restricted forms of the production function. However, these specifications can be shown to be outperformed by alternative specifications. The best overall fits are obtained in the context of both in the three and two level CES specifications where via a CES aggregator skilled and unskilled labor form a compound factor in the production function and the widening of the skill premium is explained by the markedly faster skill augmenting technical change than that of unskilled labor.

The paper is organized as follows. The next section describes data relating to the skill premium. Section 3 derives the multi-level, multi-factor CES production-technology systems. In each case (of three and two-level), we derive the first-order conditions and the skill premium in both level and growth form. We use the latter to show the decomposition of the skill wage premium into its constituent channels. Section 4 explains the US data: that relating to college-educated labor and wages, as well as the data related to output, capital inputs and relative user costs.

Section 5 then reports the estimation results. We show the estimates of the production-technology systems (the production elasticities, and the technical change growth rates). Then we apply these estimated systems for growth accounting and skill-premium accounting exercises. Section F shows how different representations of the skill premia can generate different future projections for the premium and relative prices for different scenarios relating to growth in input factors. Finally, we conclude.

2 THE SKILL PREMIUM AND ITS DETERMINANTS

Data by skill levels were obtained from Autor et al. (2008) and Acemoglu and Autor (2011). (S)killed workers are defined as those with (some) college education and above. (U)nskilled workers are defined as those with education levels up to (and including) High School. The skill premium is the difference between the wages of skilled and unskilled labor. Data is available for relative supply and relative wages for both labor types. Here relative supply is defined in terms of hours worked.[2]

Figure 1 plots various series relevant to the evolution of the skill premium over 1963-2008, and **Table 1** provides summary statistics. Section 4 explains the construction of the entire database in greater detail.

– Figure 1 here –

Panel (a) plots the log skill premium. For $\Omega = W_S/W_U$, we have $\omega = \omega_S - \omega_U$ (where $\omega_S = \ln W_S, \omega_U = \ln W_U$), and the relative supply of skills, $\frac{\mathbf{S}}{\mathbf{U}}$. This makes clear that, despite the greater supply of college-educated labor, its relative reward/price increased, revealing the hike in demand. And *Panel (d)* shows the evolution of both labor types. Table 1 further confirms that growth in skilled labor and skilled wages exceeded their unskilled counterparts: $g^S \gg g^U$, $g^{W_s} > g^{W_u}$ respectively. *Panels (e), (f)*, which demean the growth rates, shows that the pair of series display roughly similar business-cycle turning points.[3]

The figure also shows capital developments (*Panels b, c*). The (E)quipment capital to output ratio, K^e/Y, displays a positive trend and the structures (or (B)uildings) capital ratio, K^b/Y, a negative one. As Table 1 shows equipment capital relative to the structures capital rises and is reflected by the downward trend in their relative user prices. As these opposite trends largely compensate each other, their relative factor income shares remain relatively stable, only marginally favoring equipment capital. Finally, equipment capital has grown twice as fast and is twice as volatile as building capital. These features of the data suggest that the substitution elasticity between two capital inputs deviate less from unit elasticity than that between two labor inputs.

However, a deeper understanding of the interactions between all of the variables underpins the importance of applying (as we do) a *system* estimation framework with coherent cross-equation parameter restrictions. This allows us not only to assess how well we capture the skill premium but also the other variables of interest, e.g., potential output and relative user costs.

3 MULTI-LEVEL MULTI-FACTOR CES PRODUCTION FUNCTIONS

Our production technology assumption is the multi-level four factor CES production function, where the value-added of production is defined in terms of structures capital, equipment capital, skilled labor and unskilled labor. In comparison, for instance, to the more general alternative of the translog production function the indisputable advantage of the adopted CES framework is that in the latter case the number of estimated parameters is markedly fewer. That is especially the case, when the CES function is specified in the normalized form that allows to define distribution parameters as the factor income shares of the sample data. Hence, estimation is concentrated just to those parameters that are in the focus of our interest, i.e. to three substitution elasticity parameters and the four parameters of augmenting technical change. Naturally the simplicity of the framework has also its costs. That is related to the fact that in the context of the four factor CES function the production function must be defined either as a two level or as a three level function to allow for the non-equal elasticities of substitution between four factors. In the two level CES function the higher level CES is defined as the aggregate of two lower level CES aggregates each of which are formed by two inputs. In the three level CES any pair of inputs can form the lowest level CES that together with the third input forms the second level CES. Finally the highest level is formed by the fourth factor together with the combined factor determined by the second level CES aggregate. The disadvantage of this CES framework is that neither two nor three level CES nests another as a special case but each of them must be treated as separate specifications. Likewise inside both the two and three level CES specifications the way in which inputs are combined into CES sub-aggregates must be done on a priori bases. Hence, only non-nested test statistics can be applied in evaluating the data compatibility of each specification alternative.

In the following sections, we more closely outline the various hierarchies of CES production-technology systems that we estimate. In each case, we also derive the closed-form of the level and growth rate of the skill premium. The latter facilitates in a transparent way the growth decomposition of the skill wage premium (and its contributory channels), which we examine later in the paper. We also include the special case examined by Krusell et al. (2000) and demonstrate how technical progress terms may be identified in this case. Then section 3.3 makes the equivalent analysis for the four-factor, two-level CES case.

Specifically section 3.1 describes the four-factor, three-level CES system. The four (input) factors are skilled and unskilled labor and equipment capital and structures capital. The three levels points to how we combine these various input factors in producing output. If all of those factors were combined together in a single production function, then we would have a 1-level arrangement. Likewise if we had a system where the final production function subsumed two separate production functions (say 1 level for unskilled and skilled labor and one for the two capital types) then we would have a 2-level system. The advantage of going beyond a single-level production function is flexibility. The fewer levels of the production system the more you constrain the elasticities of production between factors. We also include the special case examined by Krusell et al. (2000) and demonstrate how technical progress terms may be identified in this case. Then section 3.3 makes the equivalent analysis for the four-factor, two-level CES case.

3.1 Four-Factor, Three-Level CES

Let us write the four-factor, three level-CES production function for the "normalized" production $\widetilde{Y} = Y/Y_0$ in terms of the four indexed inputs $\widetilde{V}_i = V_i/V_{i0}$ as follows (suppressing time subscripts for legibility),

$$\widetilde{Y} = \left[\alpha_0 \left(A_1 \widetilde{V}_1\right)^{\frac{\psi-1}{\psi}} + (1-\alpha_0) Z^{\frac{\psi-1}{\psi}}\right]^{\frac{\psi}{\psi-1}} \tag{1}$$

$$Z = \left[(1-\beta_0) \left(A_2 \widetilde{V}_2\right)^{\frac{\sigma-1}{\sigma}} + \beta_0 X^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \tag{2}$$

$$X = \left[(1-\pi_0) \left(A_3 \widetilde{V}_3\right)^{\frac{\eta-1}{\eta}} + \pi_0 \left(A_4 \widetilde{V}_4\right)^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}} \tag{3}$$

where ψ is the elasticity of substitution between the input V_1 and the compound input Z, σ is the parameter of the elasticity of substitution between the input V_2 and the compound input X and η is parameter of the elasticity of substitution between inputs V_3 and V_4. Parameters α_0, β_0 and π_0 are the respective factor income shares at the point of normalization (i.e., sample averages, see Appendix A). Terms A_j, $j \in [1, 4]$ represent factor-specific technical progress that are defined as the following exponential functions of time $A_j = e^{\gamma_j \tilde{t}}$ with $\tilde{t} = t - t_0$.[4]

Inserting (2) and (3) into (1), the three level-CES production function becomes,

$$\widetilde{Y} = \left[\alpha_0 \left(A_1 \widetilde{V}_1\right)^{\frac{\psi-1}{\psi}} + (1-\alpha_0) \left\{ \begin{array}{c} (1-\beta_0) \left[A_2 \widetilde{V}_2\right]^{\frac{\sigma-1}{\sigma}} + \\ \beta_0 \left[(1-\pi_0) \left(A_3 \widetilde{V}_3\right)^{\frac{\eta-1}{\eta}} + \pi_0 \left(A_4 \widetilde{V}_4\right)^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1} \frac{\sigma-1}{\sigma}} \end{array} \right\}^{\frac{\sigma}{\sigma-1} \frac{\psi-1}{\psi}}\right]^{\frac{\psi}{\psi-1}} \tag{4}$$

Assume that the representative firm faces an isoelastic demand curve, $Y_{it} = \left(\frac{P_{it}}{P_t}\right)^{-\varepsilon} Y_t$. Profit maximizing under the specified CES technology, (4), yields the four first order conditions (which for brevity are relegated to Appendix A). These equations define a 5-equation system with manifest cross-equation parameter constraints. This system encompasses the 3-equation system estimated by Krusell et al. (2000) who constrained the elasticity of substitution, ψ, between variable V_1 (structures capital) and the compound factor Z (capturing unskilled labor V_2, equipment capital V_3 and skilled labor V_4) to unity, i.e. Cobb Douglas.[5] We now more closely examine that special case.

3.1.1 Special Case: Four-Factor-Nested Cobb Douglas-CES

Under the limiting case of $\psi = 1$ we end up with two variants of the following nested Cobb Douglas-CES production function. First, the "pure" Hicks case:

$$\widetilde{Y} = A_H \widetilde{V}_1^{\alpha_0} \left[(1-\beta_0)\, \widetilde{V}_2^{\frac{\sigma-1}{\sigma}} + \beta_0 \left[(1-\pi_0)\, \widetilde{V}_3^{\frac{\eta-1}{\eta}} + \pi_0^{\frac{\eta-1}{\eta}}\, \widetilde{V}_4 \right]^{\frac{\eta}{\eta-1}\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma(1-\alpha_0)}{\sigma-1}} \tag{4'}$$

where $A_H = e^{\gamma_H \tilde{t}}$ denotes Hick-neutral technical progress and where the distribution parameters are defined by (A.5)-(A.7). As we shall see this very restrictive case (i.e., it imposes that technical progress that is common to all factors) performs (absolutely and relatively) poorly.

Moreover, form (4′) does not permit the identification of all four factor augmenting components of technical change. However, as (4″) and (5) below show, the three components of the biased technical change can be expressed in terms of the technical change of the reference factor:

$$\widetilde{Y} = \left(A_H \widetilde{V}_1\right)^{\alpha_0} \left[(1-\beta_0) \left(e^{\gamma_{24} t} \widetilde{V}_2\right)^{\frac{\sigma-1}{\sigma}} + \beta_0 \left[(1-\pi_0) \left(e^{\gamma_{34} t} \widetilde{V}_3\right)^{\frac{\eta-1}{\eta}} + \pi_0^{\frac{\eta-1}{\eta}}\, \widetilde{V}_4 \right]^{\frac{\eta}{\eta-1}\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma(1-\alpha_0)}{\sigma-1}} \tag{4''}$$

where

$$\gamma_H = \alpha_0 \gamma_1 + (1-\alpha_0)\, \gamma_4, \gamma_{24} = \gamma_2 - \gamma_4, \gamma_{34} = \gamma_3 - \gamma_4 \tag{5}$$

In this second special case, (4″), the reference factor is arbitrarily chosen to be V_4. The implied first order conditions corresponding to equations (A.1)-(A.4) are presented in Appendix B.

3.2 The Derivation of the Skill Premium

The four factor inputs are: Structures (Building) capital (K^b), Equipment capital (K^e), Unskilled labor (U) and Skilled labor (S). To proceed let us fix $V_1 = K^b$. The inputs $V_2 - V_4$ will be combinations of K^e, U and S. This implies the following three different ways (or constellations) to define the right-hand-side of equation (4) or the special cases (4′; 4″), as

$$V_1, \psi, [V_2, \sigma, (V_3, \eta, V_4)] \Leftrightarrow \begin{cases} 1. \quad K^b, \psi, \quad [S, \sigma, (K^e, \eta, U)] \\ 2. \quad K^b, \psi, \quad [U, \sigma, (K^e, \eta, S)] \\ 3. \quad K^b, \psi, \quad [K^e, \sigma, (U, \eta, S)] \end{cases} \tag{6}$$

Corresponding to constellation **1.** above, it can be shown that the skill-premium is determined by the difference of appendix equations (A.2) and (A.4):

$$\ln\left(\frac{w_2}{w_4}\right) = \mathcal{C}_2 - \mathcal{C}_4 + \left(\frac{\sigma-1}{\sigma}\gamma_2 - \frac{\eta-1}{\eta}\gamma_4\right)\tilde{t} - \frac{(\sigma-\eta)}{\eta\sigma}\ln X^{\frac{\sigma-1}{\sigma}} - \frac{1}{\sigma}\ln\widetilde{V}_2 + \frac{1}{\eta}\ln\widetilde{V}_4 \tag{7}$$

In constellation **2.** above, equation (7) defines the inverse of the skill premium. In case **3.**, the premium is defined by the difference of (A.4) and (A.3),

$$\ln\left(\frac{w_4}{w_3}\right) = \mathcal{C}_4 - \mathcal{C}_3 + \left(\frac{\eta-1}{\eta}\right)(\gamma_4-\gamma_3)\,\tilde{t} - \frac{1}{\eta}\left(\ln\widetilde{V}_4 - \ln\widetilde{V}_3\right) \tag{8}$$

where $\tilde{V}_3$ and $\tilde{V}_4$ refer to unskilled and skilled labor, respectively. In each of the three cases, the development of structures capital as well as the technical change augmenting it, have no effect on the skill premium. This is because structures capital and the compound input Z capturing the other three variables are separable.

However, in constellations 1 and 2 the skill premium is not independent from the development of equipment capital nor from the equipment capital augmenting technical change. This is because equipment capital is another component of the compound input X.

In constellation **3.** above, in turn, X is the CES index of aggregate labor input separable from equipment capital. Hence, the skill premium depends only on the relative growth rates of skilled and unskilled labor and the speeds of the technical change augmenting them. In fact, equation (8) is the specification of Katz and Murphy (1992) and Goldin and Katz (2010); a specification that Acemoglu and Autor (2011) call the "canonical" model.

To better understand the growth of the skill premium as a function of different components, we differentiate (7) and (8) at the point of normalization $t = t_0$. By denoting $g^M = \ln\left(\frac{M_{t_0+1}}{M_{t0}}\right)$ for each variable M we obtain the following relations (corresponding to the constellations in (6)):

Constellation 1:

$$\begin{aligned} g^{\omega} = & -\frac{1}{\sigma}g^S + \left[\frac{(1-\pi_0)\,\sigma + \pi_0\eta}{\eta\sigma}\right]g^U && : RSS \\ & + \frac{(1-\pi_0)\,(\eta-\sigma)}{\eta\sigma}g^{K^e} && : KSC \\ & + \frac{\sigma-1}{\sigma}\gamma^S - \left[1 - \frac{1}{\eta} - \frac{\pi_0\,(\eta-\sigma)}{\eta\sigma}\right]\gamma^U + \frac{(1-\pi_0)\,(\eta-\sigma)}{\eta\sigma}\gamma^{K^e} && : TC \end{aligned} \tag{9}$$

Constellation 2:

$$
\begin{aligned}
g^{\omega} = & -\left[\frac{(1-\pi_0)\,\sigma + \pi_0\eta}{\eta\sigma}\right] g^S + \frac{1}{\sigma} g^U && : RSS \\
& + \frac{(1-\pi_0)\,(\sigma-\eta)}{\eta\sigma} g^{K^e} && : KSC \\
& + \left[1 - \frac{1}{\sigma} - \frac{(1-\pi_0)\,(\sigma-\eta)}{\eta\sigma}\right]\gamma^S - \frac{\sigma-1}{\sigma}\gamma^U + \frac{(1-\pi_0)\,(\sigma-\eta)}{\eta\sigma}\gamma^{K^e} && : TC
\end{aligned} \tag{10}
$$

Constellation 3:

$$
\begin{aligned}
g^{\omega} = & -\frac{1}{\eta}\left(g^S - g^U\right) && : RSS \\
& + \frac{\eta-1}{\eta}\left(\gamma^S - \gamma^U\right) && : TC
\end{aligned} \tag{11}
$$

where RSS, KSC and TC respectively denote the contribution to the growth of the premium from the Relative Labor Supply effect, Capital-Skill Complementarity and Technical Change. We discuss these in turn.

3.2.1 RSS

In each case, growth of skilled (unskilled) labor decrease (increases) the premium. Since $g^S > g^U$, recall table 1, this channel has negative impact on the skill premium at least in constellation 3 whilst the magnitude of this negative effect decreases the higher is the substitution elasticity between skilled and unskilled labor η.

In constellation 1, however, the condition for a negative RSS effect is given by the inequality $g^S > \left[(1-\pi_0)\,\frac{\sigma}{\eta} + \pi_0\right] g^U$. This holds in the data, at least, if the substitution elasticity between equipment capital and skilled labor is smaller than the substitution elasticity between equipment capital and unskilled labor, i.e., $\sigma < \eta$. However in the opposite case with sufficiently high values of σ, the RSS effect may turn positive.

In constellation 2, combined with $g^S > g^U$, the condition for a negative RSS effect is $g^S > \frac{1}{(1-\pi_0)\frac{\sigma}{\eta}+\pi_0} g^U$. That inequality always holds if $\sigma > \eta$. However, symmetrically with constellation 1 the RSS effect may turn positive with low enough values of the ratio $\frac{\sigma}{\eta}$ i.e. unskilled labor is a low substitute to both equipment capital and skilled labor while equipment capital and skilled labor are high substitutes for each other.

3.2.2 KSC

In constellations 1. and 2. the growth of equipment capital affects the premium depending on its complementarity with either skilled or unskilled labor. The sign of KSC channel is positive, if equipment capital is a closer substitute to unskilled labor than to skilled labor (in constellation 1 $\eta > \sigma$; in constellation 2 $\sigma > \eta$). It is worth noticing that, if $g^S > g^U$, then whilst the widening of these substitution elasticity differences strengthen positive capital skill complementarity effect on the skill premium it also strengthens the negative RSS effect on the premium. Notice, there is no KSC in constellation 3, because skilled and unskilled labor are treated as a compound input separable from capital inputs.

3.2.3 TC

Factor augmenting technical change will affect the premium unless there are the following combinations: $\gamma^S = \gamma^U = \gamma^{K^e}$ (constellations 1. and 2.) or $\gamma^S = \gamma^U$ and/or $\eta = 1$ (constellation 3.).

In constellation 1, assuming $\left\{\gamma^S, \gamma^U, \gamma^{K^e}\right\} > 0$ positive contributions to the premia respectively requires $\sigma > 1$; $\frac{\eta\pi_0}{\eta-1+\pi_0} > \sigma$; and $\eta > \sigma$. In constellation 2. it is the reverse. And for constellation 3. we require $\eta > 1$ and $\gamma^S > \gamma^U$ or $\eta < 1$ and $\gamma^S < \gamma^U$.

3.3 Two Level CES

An alternative to the above case(s) is the four-factor-two-level CES function.[6] It contains the same number of parameters as the three level function (4). In terms of the possible range of cross-factor substitution possibilities, though, they are different. Reflecting these differences the skill premium is independent from the accumulation (and technical change) of one input (under our assumptions structures capital) in the context of the three level CES, whilst in the context of the two-level CES this is not possible. Hence, in the latter case the capital-skill complementarity channel of the skill premium is associated to *both* structures and equipment capital.

However, as neither form of the four factor CES function contains the other as a special case, there is no a priori reason to favor either. It is, however, apparent that with some appropriate combinations of the estimated parameter values the two- and three-level-CES systems may quite closely approximate each other.

The four-factor two-level-CES production function is:

$$\widetilde{Y} = \left[\alpha_0 X_1^{\frac{\sigma-1}{\sigma}} + (1-\alpha_0) X_2^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \tag{12}$$

where σ is the elasticity of substitution between compound inputs X_1 and X_2, defined by,

$$X_1 = \left[(1-\beta_0)\left(A_1\widetilde{V}_1\right)^{\frac{\zeta-1}{\zeta}} + \beta_0\left(A_2\widetilde{V}_2\right)^{\frac{\zeta-1}{\zeta}}\right]^{\frac{\zeta}{\zeta-1}} \tag{13}$$

$$X_2 = \left[(1-\pi_0)\left(A_3\widetilde{V}_3\right)^{\frac{\eta-1}{\eta}} + \pi_0\left(A_4\widetilde{V}_4\right)^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}} \tag{14}$$

where ζ and η are the respective elasticity of substitutions between inputs V_1 and V_2, and between V_3 and V_4. Inserting (13) and (14) into (12) the two-level-CES production function becomes,

$$\widetilde{Y} = \left[\begin{array}{r} \alpha_0\left[(1-\beta_0)\left(A_1\widetilde{V}_1\right)^{\frac{\zeta-1}{\zeta}} + \beta_0\left(A_2\widetilde{V}_2\right)^{\frac{\zeta-1}{\zeta}}\right]^{\frac{\zeta}{\zeta-1}\frac{\sigma-1}{\sigma}} + \\ (1-\alpha_0)\left[(1-\pi_0)\left(A_3\widetilde{V}_3\right)^{\frac{\eta-1}{\eta}} + \pi_0\left(A_4\widetilde{V}_4\right)^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}\frac{\sigma-1}{\sigma}} \end{array}\right]^{\frac{\sigma}{\sigma-1}} \tag{15}$$

We relegate the first order conditions to Appendix C.

3.3.1 Skill premium

Within this setting we shall derive different skill-premium relations for each of the three alternative ways to combine inputs in the two-level CES,

$$(V_1,\zeta,V_2),\sigma,(V_3,\eta,V_4) \Leftrightarrow \begin{cases} 1. \quad (K^b,\zeta,S) \quad ,\sigma, \quad (K^e,\eta,U) \\ 2. \quad (K^b,\zeta,U) \quad ,\sigma, \quad (K^e,\eta,S) \\ 3. \quad (K^b,\zeta,K^e) \quad ,\sigma, \quad (U,\eta,S) \end{cases} \tag{16}$$

Constellation 1:

$$\begin{aligned} g^{\omega} = &-\left[\frac{1-\beta_0}{\zeta}+\frac{\beta_0}{\sigma}\right] g^S + \left[\frac{1-\pi_0}{\eta}+\frac{\pi_0}{\sigma}\right] g^U && : RSS \\ &+ (1-\pi_0)\,\frac{\eta-\sigma}{\eta\sigma} g^{K^e} + (1-\beta_0)\,\frac{\sigma-\zeta}{\sigma\zeta} g^{K^b} && : KSC \\ &\left.\begin{aligned} &+\left(1-\left[\tfrac{1-\beta_0}{\zeta}+\tfrac{\beta_0}{\sigma}\right]\right)\gamma^S - \left(1-\left[\tfrac{1-\pi_0}{\eta}+\tfrac{\pi_0}{\sigma}\right]\right)\gamma^U \\ &+ (1-\pi_0)\,\tfrac{\eta-\sigma}{\eta\sigma}\gamma^{K^e} + (1-\beta_0)\,\tfrac{\sigma-\zeta}{\sigma\zeta}\gamma^{K^b} \end{aligned}\right\} && : TC \end{aligned} \tag{17}$$

Constellation 2:

$$\begin{aligned} g^{\omega} = &-\left[\frac{1-\pi_0}{\eta}+\frac{\pi_0}{\sigma}\right] g^S + \left[\frac{1-\beta_0}{\zeta}+\frac{\beta_0}{\sigma}\right] g^U && : RSS \\ &- (1-\pi_0)\,\frac{\eta-\sigma}{\eta\sigma} g^{K^e} - (1-\beta_0)\,\frac{\sigma-\zeta}{\sigma\zeta} g^{K^b} && : KSC \\ &\left.\begin{aligned} &+\left(1-\tfrac{1-\pi_0}{\eta}-\tfrac{\pi_0}{\sigma}\right)\gamma^S - \left(1-\tfrac{1-\beta_0}{\zeta}-\tfrac{\beta_0}{\sigma}\right)\gamma^U \\ &- (1-\pi_0)\,\tfrac{\eta-\sigma}{\eta\sigma}\gamma^{K^e} - (1-\beta_0)\,\tfrac{\sigma-\zeta}{\sigma\zeta}\gamma^{K^b} \end{aligned}\right\} && : TC \end{aligned} \tag{18}$$

Constellation 3:

$$\begin{aligned} g^{\omega} = &-\frac{1}{\eta}\left(g^S - g^U\right) && : RSS \\ &+\frac{\eta-1}{\eta}\left(\gamma^S-\gamma^U\right) && : TC \end{aligned} \tag{19}$$

3.3.2 RSS

In constellation 1. the RSS channel is positive if $\frac{g^U}{g^S} > -\frac{\frac{1-\pi_0}{\eta}+\frac{\pi_0}{\sigma}}{\frac{1-\beta_0}{\zeta}+\frac{\beta_0}{\sigma}}$, the reverse in constellation 2. In constellation 3., the effect is negative for $\eta > 0$.

3.3.3 KSC

A key implication of KSC is that growth in the stock of equipment capital increases [decreases] the marginal product and the wage rate of skilled [unskilled] labor. This is so if $\eta > \sigma$, i.e. unskilled labor is a closer substitute to equipment capital than to the composed input of skilled labor and structures capital. We also find that the growth in the stock of structures capital has similar effect, if $\zeta < \sigma$ i.e. skilled labor and structures capital are weaker substitutes to each other than to two other inputs. Hence, under the ordering $\eta > \sigma > \zeta$ the growth of both capital inputs expand the premium in constellation 1. This is the main qualitative difference compared to the three level cases where the skill premium was unaffected by structures capital. We see, however, that this ordering tends to also strengthen the negative RSS effect related to the faster growth of skilled than unskilled labor compared to e.g. the opposite ordering. In constellation 2. the order reverses: $\zeta > \sigma > \eta$. Constellation 3. echoes (11) with KSC playing no role.

3.3.4 TC

Constellation 1. requires $\frac{1-\left[\frac{1-\beta_0}{\zeta}+\frac{\beta_0}{\sigma}\right]}{1-\left[\frac{1-\pi_0}{\eta}+\frac{\pi_0}{\sigma}\right]} > \frac{\gamma^U}{\gamma^S}$ to generate a positive contribution to the skill premium from the labor-augmenting terms, and $\eta > \sigma > \zeta$ (assuming $\{\gamma^{K^e}, \gamma^{K^b}\} > 0$). Constellation 2 is the reverse. In constellation 3. the TC channel is positive if labor inputs are gross substitutes [complements] and $\gamma^S > \gamma^U$ [$\gamma^S < \gamma^U$]. We now discuss the data as a prelude to presenting the empirical estimates for our various specifications.

4 DATA

Annual data were obtained from various sources for the US economy for period 1963-2008. The frequency is determined by the availability of skilled/unskilled hours and wages. Data for output, capital, total employment, and labor compensation are for the US private non-residential sector. Most of the data come from NIPA series available from the Bureau of Economic Analysis (BEA). The output series are thus calculated as total output minus net indirect tax revenues, public-sector, and residential output. After these adjustments, the output concept used is compatible with that of the private non-residential capital stock (chain-type quantity indexes and current cost net stocks) and depreciation series obtained from BEA's fixed assets accounts (tables 2.1, 2.2 and 2.4). The base year in all constant dollar series is 2005. Non-residential capital stock and related depreciation data is disaggregated into non-residential equipment and structures data. Data by skill levels were obtained from Autor et al. (2008) and Acemoglu and Autor (2011). Skilled workers are defined as those with (some) college education and above. Unskilled workers are defined as those with education levels up to (and including) High School. Autor et al. (2008) provide relative supply and relative wages for both categories. Relative supply is defined in terms of hours worked.[7]

Because the coverage of these data coming from the Current Population Survey is different from our coverage for the non-residential private sector, we combined these data with Bureau of Labor Statistics

(BLS) data. While preserving relative wages and relative labor supply, we correct both so as to be compatible with the evolution of total private employment and labor compensation. Hence, we proceed as follows. We define unskilled workers' wages (W^U) as, $W^U = \frac{W}{U/N+(S/N)\times\Omega}$ where W are wages of all workers, $N = S + U$ is total private sector workers. Variable Ω, as before, is the level skill premium, $\Omega = W_S/W_U$. Then skilled wages are defined as $W^S = W \times \Omega$.

We now need to define how some of these variables are obtained. We define W as labor income (INC_N) over total private sector employment. A problem in calculating labor-income is that it is unclear how the income of proprietors (self-employed) should be categorized in the labor-capital dichotomy. Some of the income earned by self-employed workers clearly represents labor income, while some represents a return on investment or economic profit. Following Blanchard (1997), Gollin (2002), Klump et al. (2007), and McAdam and Willman (2013) we use compensation per employee as a shadow price of labor of self-employed workers:[8] $INC_N = \left(1 + \frac{\text{self-employed}}{\text{total private employment}}\right) \cdot \text{Comp}$ where $Comp$ = private sector compensation of employees.

We then define $W = \frac{INC_N}{\text{total private sector employment}}$. Finally, we define S as total private sector employment times relative skilled/unskilled hours worked, and $U = N - S$. These transformations preserve relative quantities but correct the levels in order to comply with our previous definitions and the self-employment transformation. This assumes, of course, that relative wages and relative labor supply in the private sector evolve in a similar fashion to those in the (wider) definition provided by Autor et al. (2008).

Total capital income was obtained using a residual approach:$INC_K = \frac{Y \cdot P}{1+\mu} - INC_N$ where Y is the real GDP of the private non-residential sector, P is the GDP deflator, and the aggregate mark-up was set to $\mu = 0.1$ (in line with empirical estimates, e.g., Klump et al. (2007)).[9]

NIPA does not offer a direct way to disaggregate INC_K into capital income related to equipment on one hand and structures on the other hand. Therefore, we use a two-step approach. In the first step we construct the direct estimates of the two capital income components as follows, $INC_{K^e} = \left(\frac{i}{100} - \overline{\pi}^e - \delta_t^e\right) K_t^e \cdot P_t^{K^e}$ and $INC_{K^b} = \left(\frac{i}{100} - \overline{\pi}^b - \delta_t^b\right) K_t^b \cdot P_t^{K^b}$ where i is the ten-year treasury bond rate; $\overline{\pi}^e$ and $\overline{\pi}^b$ are the sample averages of the inflation rates of the respective investment deflators; δ_t^e and δ_t^b are the depreciation rates (calculated as ratios of current-cost depreciations to the two year end-of-year current-cost net stocks); K_t^e and K_t^b are the two-year end averages of year-end constant dollar net capital stocks; and, finally, $P_t^{K^e}$ and $P_t^{K^b}$ are the implied deflators of capital stocks. Naturally the summation, $SUM = INC_{K^e} + INC_{K^b}$ does not coincide with INC_K but their general developments are very similar. By assuming that income shares in terms of INC_K follow those defined by INC_{K^e} and INC_{K^b}, the real user cost developments of equipment capital and structures capital in terms of NIPA data are defined as, $uc^e = \frac{\frac{INC_{K^e}}{SUM} INC_K}{K^e P}$ and $uc^b = \frac{\frac{INC_{K^b}}{SUM} INC_K}{K^b P}$

5 ESTIMATION RESULTS

5.1 Metrics and Estimations

Our estimation results report the various substitution elasticities, ψ, σ, η and ζ; and the growth in factor-augmenting technical progress components, $\gamma^U, \gamma^S, \gamma^{K^b}, \gamma^{K^e}$ and in the special "pure" Hicks case, γ^H. Below these are various restrictions: tests of unitary and common elasticities, pairwise-equality in factor-augmenting technical progress; a test for Hicks Neutrality, $\gamma^i = \gamma^j, \forall \{i, j\}$ and p(robability)-values from ADF residuals tests (where the null is non-stationarity). In most cases, these parameter restrictions are not accepted by the data; where accepted, we additionally impose them for comparability, in subsequent columns. After the main results, we present a measure of the fit of the estimated system of equations, the determinant of the residual covariance, $|VC_\varepsilon|$ and where useful, we discuss the residual properties of the residuals of the system equations. In Appendix E we also present the residual standard errors of each equation of estimated systems for highlighting which equations contribute most to the overall fit.[10] Finally, note we conducted extensive robustness and sensitivity checks. Initial conditions of all parameters were varied around plausible supports to ensure a global optimum. Our estimation method is non-linear, iterative Seemingly Unrelated Regression (SUR).[11] This estimates the non-linear system accounting for heteroskedasticity and contemporaneous correlation in the errors across the equations. In estimation, we also respect cross-equation parameter restrictions. The estimation method is simultaneously iterative in terms of the weighting matrix in the covariance calculation, and the model parameters until full convergence.

5.2 Decompositions

In addition to the main results, we derive the contributions of each component on the skill premium and on output growth. The contribution of each component is calculated as the difference of the estimated equation for the skill premium (or output) when one component at a time gets first current and then one period lagged values whilst all other variables get current period values. The method is explained in full in Appendix D. As before the skill-premium decomposition is given by,$\widehat{g^{\omega}} = \omega^{RSS} + \omega^{KSC} + \omega^{TC}$ where, in line with our previous classifications, $\omega^{RSS} = \omega^S + \omega^U, \omega^{KSC} = \omega^{K^e} + \omega^{K^b}$ and, $\omega^{TC} = \omega^{TC^S} + \omega^{TC^U} + \omega^{TC^{K^e}} + \omega^{TC^{K^b}}$. For output growth, we have as normal, $\widehat{g^Y} = y^N + y^K + y^{TFP}$ where, $y^N = y^U + y^S, y^K = y^{K_b} + y^{K_e}$ and, $y^{TFP} = y^{TC^S} + y^{TC^U} + y^{TC^{K^e}} + y^{TC^{K^b}}$ The terms $\widehat{g^{\omega}}$ and $\widehat{g^Y}$ can then be compared with their observed counterparts.

5.3 Estimation Results: Three-Level CES

Tables 2-4 presents the 3-level cases in the order indicated in (6). Columns (a) presents the estimation results of the most general specifications. Column (b) in each table presents estimation results – corresponding to Krusell et al. (2000) – under a unit elasticity constraint $\psi = 1$ and "pure" Hicks neutral technical change, as in $(4')$. In line with $(4'')$ columns (c) of the tables allow a more general factor augmenting technical change in the $\psi = 1$ case. We see that estimation results under Hicks neutrality condition (columns (b)) are extremely poor in terms of fit;[12] plus the residual properties suggest unit roots for 3-4 residuals out of 5 in all cases. This is the case over all the estimation tables. Hence, the allowance of a more general factor augmenting technical change is crucial and in the following we concentrate columns (a) and (c) of tables 2-4. (see also our later footnote 15.)

– Tables 2, 3, and 4 here –

To recall, the first specification (Table 2) implies KSC if $\eta > \sigma$. In the second, KSC implies the reverse ordering: $\eta < \sigma$ (Table 3). The third (Table 4) does not admit capital-skill complementarity. **Tables 5-6** show the decomposition of the skill premium into its standard parts, as well as growth accounting disaggregating by factor accumulation and TFP growth (with individual components therein). The growth accounting exercises show a relatively robust division of the drivers of growth across specifications. For example, with labor accumulation accounting for just under half of growth and capital and TFP split roughly equally of the remainder.

– Tables 5 and 6 here –

5.4 Constellation 1

Specification alternatives presented in Table 2 allows KSC, if substitution elasticity $\eta > \sigma$. However, all columns of Table 2 indicate just the opposite. Hence, under this specification alternative (positive) KSC does not hold. For instance, the most general specification alternative that is presented in column (a) indicates a high substitution elasticity between skilled labor and equipment capital ($\sigma = 4.03$), well above the elasticity of equipment capital to unskilled labor ($\eta = 0.66$). All parameter estimates are statistically significant and the constraint $\psi = 1$ is rejected at 1%. Diagnostics reveals only that a common technical progress growth rate for skilled and unskilled labor cannot be rejected at 15% and, except the inclusion of this parameter constraint, column (a') basically repeats the results of column (a). However, the comparison of results presented in columns (a) and (a') to those in column (c) that is estimated under the unit substitution elasticity constraint $\psi = 1$ show that the latter alternative is markedly better indicating smaller residuals (see VC statistics) as well as better stationarity properties of residuals.

A plausible explanation to the counterintuitive result that the specification containing parameter constraints give the better fit than the unconstrained specification is that the unconstrained estimation results in column (a) represent local instead of global optimum. Therefore, we experimented on alternatives where the parameters of augmented technical change contained constraints in a nonstandard way whilst substitution elasticity parameters are freely estimated. The best estimation results in that regard are presented in column (a''), where the values of equipment and unskilled labor augmenting technical change are constrained to be equal although with opposite signs.[13] The fit is relatively good and the ADF t-test indicated below the 5% probability of unit root for all residuals. Although these results indicates quite a high substitution elasticity between equipment capital and unskilled labor ($\sigma = 1.63$) the KSC hypothesis is not supported since skilled labor is even closer substitute to equipment capital ($\eta = 2.43$). The first column of Table 5 confirms this implying negative impact of both RSS and KSC channels on the skill premium.

Accordingly, under this specification it is technical change that explains the widening skill premium. A more detailed examination reveals that it is mainly the skill augmenting technical change difference relative to that of unskilled labor with a minor negative impact from equipment augmenting technical change that explains the widening skill premium under this specification. In terms of the growth contributions, see Table 6, all factors naturally contribute positively to growth, as does TFP (the summation of the technical change terms).[14]

5.5 Constellation 2

The second specification alternative is presented in Table 3. This alternative – under the constraint $\psi = 1$ – was supported by the estimation results of Krusell et al. (2000). *They* found (positive) capital skill complementarity in the determination of the premium; $\eta = 0.67 < \sigma = 1.67$. Our results in Table 3 are in line with theirs indicating $\eta = 0.78$ and $\sigma = 1.82$ in the unconstrained specification presented in column (a). Also the results of column (c), although the unit elasticity constraint $\psi = 1$ is not validated, are very similar.

As Table 5 confirms, there is positive KSC effect through which the fast accumulation of equipment capital is transmitted into the skill premium. This effect is strengthened by the high estimate of equipment capital augmenting technical change. With the negative contribution of RSS on the skill premium, broadly neutralized by positive technical change effect, the KSC effect is sufficient to explain the observed widening of the skill premium. Hence, as regards the KSC hypothesis these results are just the opposite to those from the first specification. However, although the results of Table 3 look in economic terms quite reasonable, the fit in terms of residual covariance determinants and stationarity test statistics are worse than in columns (a'') and (c) of Table 2. In spite of big differences in parameter estimates in the columns of Tables 2 and 3 the decompositions of growth contribution are very similar in all columns (Table 6).

5.6 Constellation 3

This specification, which does not allow KSC, gives statistically the best results and is economically quite reasonable (Table 4). The residual determinant covariance corresponding to column (a) that gives the best estimation results represents a considerable gain over the results in Tables 2 and 3. Also the unit root of residual is rejected for all equations of the system at 5% significance level and for 4 equations out of 5 at 1% level. According to these results skilled and unskilled labor are quite high substitutes ($\eta = 2.34$) with each other and, via the CES aggregator, labor can be treated a compound factor separable from capital.

The substitution elasticities of this compound labor input between both equipment and structures capital are both below unity (around 0.6). Skilled labor-augmenting technical change is clearly the dominating component of technical change. When the unskilled labor augmenting technical change is slightly negative, the growth differences of the skilled and unskilled labor augmenting technical change more than compensates the RSS effect and explains the widened skill premium (see Table 5). Technical change also augments structures capital, whilst equipment capital augmenting technical change is negative.

Hence, although technical change may be largely embodied in equipment capital, this need not imply it being equipment capital saving, especially, when the fast growth of equipment capital stock is coupled with its decreasing price. Again the growth contributions of inputs and technical changes do not differ much from those implied other specifications. Perhaps the most noticeable difference is somewhat stronger contribution of skill augmenting technical change than under other specification alternatives and minor negative contributions of both unskilled labor and equipment capital augmenting technical change.

5.7 Estimation results: Two-Level CES

Table 7 presents the 2-level cases (**Tables 8-9** show the decompositions). As with the 3-level case, the first specification treats K^e and U, the second K^e and S and the third S and U as a compound factor. As regards KSC results of the first 2-level specification presented in column (a) of the Table 7 correspond qualitatively those in column (a'') of Table 2. Substitutability between equipment capital and unskilled labor is quite high ($\eta = 1.70$) – very close to correspondent three level estimate – but KSC is not supported as skilled labor is even closer substitute to equipment capital ($\sigma = 2.95$). The main difference between the 2-level and the 3-level cases is that the former implies high and the latter low substitution elasticities between structures capital and other inputs.

– Tables 7, 8 and 9 here –

Regarding technical change the estimates of the structures capital augmenting and the equipment capital augmenting technical change are very different. Perhaps counter-intuitively, the 2-level specification implies a strongly negative whilst 3-level a positive estimate for the structures capital augmenting technical change. Overall fits and residual properties are quite comparable favoring, however, somewhat the 3-level specification in column (a'') in table 2. Also in terms of economic interpretation, especially regarding the capital-augmenting technical change estimates, 3-level specification looks more plausible.

The second specification alternative is presented in columns (b) of Table 7.[15] As regards KSC these estimation results are also similar to those of the corresponding 3-level specification presented in column (a) of Table 3. Equipment capital is a closer substitute to unskilled than to skilled labor supporting KSC. In addition, as structures capital is closer substitute to unskilled than skilled labor that implies KSC also between these two factors. The fit of this specification to data is better than that of corresponding 3-level specification but worse than that of the 2-level system based on specification 1 in column (a) of table 7.

As in 3-level estimations, the third specification alternative gives the statistically best results also in the context of 2-level estimation. The 2-level estimation results are presented in column (c) of Table 7. Although the 3-level results are statistically somewhat better the qualitative implications are quite similar in both cases. Estimated substitution elasticities between skilled and unskilled labor are above 2 under both specifications and the substitution elasticity between labor and both capital inputs are below 1. Also estimates of factor augmenting technical change are quite similar with skilled augmenting technical change dominating in both cases.

5.8 Comparison of estimation results

Our results show that quite reasonable estimation results are possible to obtain with several specification alternatives based on 3-level and 2-level CES production functions each of which contain different a priori constraints on how inputs are compounded. We also find that the estimation results implied by the different specification alternatives give widely differing substitutability elasticities between the four inputs and, hence, different explanations for the skill premium. Therefore we have also tried to econometrically discriminate between them.[16] The test statistics measuring the overall fit and stationarity properties of the estimation residuals of the 5-equation system are the best under 3-level CES specification where, via a CES aggregator, skilled and unskilled labor form a compound labor input with a quite high substitutability 2.3 between them (Table 4, column (a)).)

In addition, these results are economically well interpretable, which two factor production functions with capital and labor as two aggregate inputs may be envisaged to approximate. Also its counterpart, the specification with the 2-level CES, although statistically somewhat worse, imply quite similar results and is the best across three 2-level CES specification alternatives (Table 7, column (c)).

Since the specification in these two alternatives does not allow KSC, implying the "canonical" model of Acemoglu and Autor (2011) for the skill premium, the widening skill premium is explained by technical change augmenting skilled labor more than unskilled labor. We also found that both 3-level 2-level CES where, as in Krusell et al. (2000), skilled labor and equipment capital is treated, via the CES aggregator, a compounded input give reasonable and, especially, in the latter case also statistically quite satisfactory results which support KSC.

However, these results are surpassed by both other specifications and interestingly 2-level and 3-level CES specifications with equipment capital and unskilled labor treated as a compound input suggest gross substitutability between capital and skilled labor, i.e. negative KSC. Interestingly, the earlier literature has typically thought this alternative a less plausible specification.

Unlike the very different interpretations regarding the sources of the widened skill premium, it is striking to observe how similar growth contribution estimates each specification alternative gives. In all cases labor, capital and TFP have corresponded to around 47%, 21-22% and 31-32%, respectively, of the growth in the sample period. Around 89% of labor contribution and around 82% of capital contribution - except specification 1 of 2-level case where the latter contribution is 92% – is related to the growth of skilled labor and the growth of equipment capital, respectively. Only the TFP contributions varies somewhat across cases. However, the growth contribution of skilled labor augmenting technical change dominates quite uniformly other forms of technical change.

The similarity of growth contribution can be thought of as an embodiment of the Diamond-McFadden Impossibility theorem (Diamond et al. (1978)) that states that the free estimation of both the parameters of substitution elasticity and the components of augmenting technical change is impossible only from the production function. Our results, in turn, show that very different combinations of substitution elasticity and augmenting technical change estimates, based however on 5-equation systems, produce very similar growth contributions of underlying factors.

6 CONCLUSIONS

We tried, through multiple lenses, to understand the skill wage premium. This has almost doubled since the 1960s alongside a rise in the ratio of skilled-to-unskilled workers. There are two main explanations for this rising premium: capital-skill complementarity and technical changes that favored skilled over unskilled labor. Of course, many other channels affect the skill premium and have been discussed in the literature. In our framework, though, these other channels act through factor demands and technical progress.

We estimated aggregate production and factor demands (skilled/unskilled labor, and equipment and structures capital) in a manner that encompasses both main explanations and allows the data to discriminate between them. We also decomposed growth in the skill premium and to aggregate growth into three parts: one related to the relative skill effect, the second to capital-skill complementarity and the final one to technical change.

Thus, in contrast to much of the tone of the literature, we do not see these explanations as necessarily exclusive. We estimated under three level and two level production systems. And in each case, we employed several different ways of combing the nested production functions. We estimated these as a system with non-linear SUR and with cross-equation constraints respected. We also estimated under system normalization which is an essential ingredient for robust parameter identification. In short, we believe this paper to be the most rigorous and encompassing attempt so far at identifying the channels behind skill-biased technical change in a production-based framework.

In general, we found only weak evidence for capital-skill complementarity in explaining the widening skill premium. However, our results show that in isolation, one can obtain quite reasonable looking estimation results with several specification alternatives based on 3-level and 2-level CES production functions each of which contain different a priori constraints on how inputs are compounded. We also found that the estimation results implied by different specification alternatives give widely differing substitutability elasticities between the four inputs and, hence, different explanations for the skill premium.

Accordingly, when the form of the production function is appropriately constrained, results also supporting the capital skill complementarity hypothesis can be estimated in our data. However, the results based on production functions, where inputs are combined in other ways, statistically outperform those results. The best overall fits are obtained in the context of both in the 2-level and 3-level CES specifications where via a CES aggregator skilled and unskilled labor form a compound factor in the production function and the widening of the skill premium is explained by the markedly faster skill-augmenting technical change relative to that of unskilled labor.

According to those estimation results, the substitution elasticity between skilled and unskilled labor is slightly above 2 whilst that between structures and equipment capital is below 1. Likewise the elasticity of substitution between the compound labor and capital inputs are below 1. In general estimation results based on 2-level and 3-level CES cases are quite supportive of one another, whilst in our data the latter specification gave statistically the best results.

Notes

[1]Of course, many other channels affect the skill premium and have been discussed in the literature. In our framework, though, these other channels act through factor demands and technical progress. As Acemoglue (2011) states "... There is a debate in labor economics [as to] what is the role of technology, trade; but most economists are comfortable in thinking technology has played a leading role here; trade has probably played quite a major role, too, but intermediated by technology ... off-shoring and out-sourcing of jobs that has been intermediated by information technology of the late 1990s and 2000s.", see http://www.econtalk.org/archives/2011/02/acemoglu_on_ine.html

[2]See Autor et al. (2008) for further detail on data construction.

[3]Thus, developments in both relative inputs (S/U) as well as in the premium favor skilled labor, i.e. both have an upward trend implying an even steeper trend in the skilled labor income to unskilled labor income ratio. This provides indication *against* a unit substitution elasticity between these two labor inputs, since under Cobb Douglas factor shares would be constant.

[4]Normalization essentially implies representing the production function and factor demands in consistent indexed number form. It is expressed in this way since its parameters then have a direct economic and econometrically identifiable interpretation. Otherwise the estimated parameters can be shown to be scale dependent, arbitrary and unrobust. Subscripts "0" denote the specific normalization points: geometric (arithmetic) averages for non-stationary (stationary) variables. See Klump et al. (2012) for a survey, León-Ledesma et al. (2015, 2010) for a Monte-Carlo analysis and de La Grandville (1989) and Klump and de La Grandville (2000) for seminal contributions. It is straightforward to see that at the point of normalization, $t = t_0$, $\tilde{t} = 0$ and $\widetilde{Y} = \widetilde{V}_j = Z = X = A_j = 1$.

[5]This assumption allowed the dropping of the first-order condition with respect to structures capital from the system. On top of that Krusell et al. (2000) did not include the production function in their estimated system.

[6]For the theoretical foundations of the two-level CES see Sato (1967).

[7]See Autor et al. (2008) for further detail on data construction.

[8]See Mućk et al. (2015) for a review of different US labor share definitions.

[9]Attempts to estimate the mark-up as a parameter in the rest of the system did not change the results (details available). The same can be said for possible time variation in the markup since in our framework this simply shows as autocorrelated errors.

[10]Quite uniformly the residuals show that the fit of production function is the least sensitive with respect to specification alternatives, and, in absolute terms, are the smallest. Typically also the FOCs of labor inputs have smaller residual squared errors than those of the capital inputs. Across columns there are quite large differences of residual standard errors varying (at least broadly well) in line with the determinants of residual covariance.

[11]We also applied several other system estimation methods and in general they gave qualitatively and, especially, in the case of Full information Maximum Likelihood Method also numerically very similar results.

[12]A manifestation of which being the ballooning of σ up to almost 200 in the 3-level, second constellation case – albeit the value of which is not statistically significant.

[13] The estimation results of column (a'') are almost identical with those that full-information-maximum likelihood method (FIML) - containing no parameter constraints – produced. In general our experience was that, whenever – beginning from same initial values – iterative SUR was able to converge, parameter estimates were almost identical with those of FIML.

[14]Note also that although equipment-capital augmenting technical change does not impact the skill premium, it does affect growth accounting.

[15]It is worth noting that although substitution elasticity estimates of ζ are high, they are not statistically significant. This reflect the problem that it is difficult for estimation algorithms to estimate precisely very high values of substitution elasticities. We estimated equations presented in columns (b) also by specifying ζ in inverse form. Estimates of the inverse ζ were close to zero and not statistically different from zero supporting the constraint $\zeta = \infty$.

[16]Section F shows how different representations of the skill premia can generate different future projections for the premium and relative prices for different scenarios relating to growth in input factors.

References

Acemoglu, D. (2002). Technical change, inequality, and the labor market. *Journal of Economic Literature*, 40(1):7–72.

Acemoglu, D. and Autor, D. H. (2011). Skills, tasks and technologies. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics Vol. 4*, chapter 12, pages 1043–1171. Elsevier.

Autor, D. H., Katz, L. F., and Kearney, M. S. (2008). Trends in U.S. Wage Inequality: Revising the Revisionists. *Review of Economics and Statistics*, 90(2):300–323.

Autor, D. H., Katz, L. F., and Krueger, A. B. (1998). Computing Inequality: Have Computers Changed The Labor Market? *Quarterly Journal of Economics*, 113(4):1169–1213.

Blanchard, O. J. (1997). The Medium Run. *Brookings Papers on Economic Activity*, 2:89–158.

Bound, J. and Johnson, G. (1992). Changes in the structure of wages in the 1980's: An evaluation of alternative explanations. *American Economic Review*, 82(3):371–92.

de La Grandville, O. (1989). In Quest of the Slutsky Diamond. *American Economic Review*, 79(3):468–481.

Diamond, P. A., McFadden, D., and Rodriguez, M. (1978). Measurement of the elasticity of substitution and bias of technical change. In Fuss, M. and McFadden, D., editors, *Production Economics, Vol. 2*, pages 125–147. Amsterdam and North Holland.

Dobbs, R. et al. (2012). *The World at Work: Jobs, Pay and Skills for 3.5 Billion People*. McKinsey Global Institute.

Goldin, C. and Katz, L. F. (2010). *The Race between Education and Technology*. Cambridge: Harvard University Press.

Gollin, D. (2002). Getting Income Shares Right. *Journal of Political Economy*, 110:458–474.

Gordon, R. (1990). *The Measurement of Durable Goods Prices*. University of Chicago Press.

Greenwood, J., Hercowitz, Z., and Krusell, P. (1997). Long-run implications of investment-specific technological change. *American Economic Review*, 87(3):342–62.

Griliches, Z. (1969). Capital-skill complementarity. *Review of Economics and Statistics*, 51(4):465468.

Katz, L. and Murphy, K. (1992). Changes in relative wages, 1963-1987: Supply and demand factors. *Quarterly Journal of Economics*, 107:35–78.

Klump, R. and de La Grandville, O. (2000). Economic growth and the elasticity of substitution: Two theorems and some suggestions. *American Economic Review*, 90(1):282–291.

Klump, R., McAdam, P., and Willman, A. (2007). Factor Substitution and Factor Augmenting Technical Progress in the US. *Review of Economics and Statistics*, 89(1):183–92.

Klump, R., McAdam, P., and Willman, A. (2012). The Normalized CES Production Function: Theory and Empirics. *Journal of Economic Surveys*, 26(5):769–799.

Krusell, P., Ohanian, L., Rios-Rull, J. V., and Violante, G. (2000). Capital-Skill Complementarity and Inequality. *Econometrica*, 68(5):1029–1053.

León-Ledesma, M. A., McAdam, P., and Willman, A. (2010). Identifying the Elasticity of Substitution with Biased Technical Change. *American Economic Review*, 100(4):1330–1357.

León-Ledesma, M. A., McAdam, P., and Willman, A. (2015). Production technology estimates and balanced growth. *Oxford Bulletin of Economics and Statistics*, 77(1):40–65.

McAdam, P. and Willman, A. (2013). Medium Run Redux. *Macroeconomic Dynamics*, 17(4):695–727.

Mućk, J., McAdam, P., and Growiec, J. (2015). Will the True Labor Share Stand UP? Analyzing Alternate US Labor Shares. Working Paper Series 1175, European Central Bank.

Sato, K. (1967). A Two-Level Constant-Elasticity-of-Substitution Production Function. *Review of Economic Studies*, 34(2):201–218.

Tinbergen, J. (1974). Substitution of Graduate by other Labour. *Kyklos*, 27(2):217–226.

Table 1: The Skill Premium and Determinants (Descriptive Statistics)

	g^{ω}	g^{W_s}	g^{W_u}	g^{S}	g^{U}	g^{K^b}	g^{K^e}	$\ln r^{K^b}$	$\ln r^{K^e}$	g^{Y}
Mean	0.63	1.72	1.08	3.37	0.45	2.42	4.47	−3.25	−1.33	3.53
St Dev	1.47	1.67	1.66	1.99	2.45	0.83	1.63	0.17	0.13	2.82

Note: For $x = \ln(X), g^X = x - x(-1)$. Figures scaled by 100.

Table 2: Three Level Case, Constellation 1. in equation (6).

Specification:	$K^b, \psi, [S, \sigma, (K^e, \eta, U)]$				
Case:	(a)	(a')	(a'')	(b)	(c)
ψ	0.785^{***}	0.797^{***}	0.697^{***}	1	1
σ	4.031^{***}	5.541^{***}	2.426^{***}	0.419^{***}	2.661^{***}
η	0.661^{***}	0.671^{***}	1.627^{***}	0.328^{***}	1.804^{***}
γ^{K^b}	0.049^{***}	0.050^{***}	0.016^{***}	–	–
γ^{K^e}	-0.051^{***}	-0.051^{***}	0.011^{***}	–	–
γ^S	0.018^{***}	0.016^{***}	0.026^{***}	–	–
γ^U	0.014^{***}	0.016^{***}	-0.011^{***}	–	–
γ^H	–	–	–	0.011^{***}	0.024^{***}
$\gamma^{K^e} - \gamma^S$	–	–	–	–	-0.014^{***}
$\gamma^U - \gamma^S$	–	–	–	–	-0.033^{***}
	Restrictions and Diagnostics				
$\psi = 1$	[0.000]	[0.000]	[0.000]	–	–
$\sigma = 1$	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
$\eta = 1$	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
$\sigma = \eta$	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
$\gamma^i = \gamma^j \; \forall i, j$	[0.000]	[0.000]	[0.000]	–	–
$\gamma^{K^b} = \gamma^{K^e}$	[0.000]	[0.000]	[0.000]	–	–
$\gamma^U = \gamma^S$	[0.157]	–	[0.000]	–	–
	ADF tests				
FOC_{K^b}	[0.900]	[0.884]	[0.007]	[0.027]	[0.027]
FOC_{K^e}	[0.474]	[0.423]	[0.024]	[0.520]	[0.011]
FOC_S	[0.639]	[0.621]	[0.001]	[0.233]	0.000]
FOC_U	[0.146]	[0.175]	[0.015]	[0.966]	[0.105]
CES	[0.827]	[0.859]	[0.004	[0.195]	[0.008]
$\|VC_\varepsilon\|$	12.700	13.200	2.13	843.000	3.040

Notes: ***,** and * respectively indicate the 1%, 5% and 10% level of significance. Probability values are in squared parenthesis. "—" indicates not applicable. Normalization parameters: $[\alpha_0, \beta_0, \pi_0] = [0.056, 0.509, 0.706]$. The $|VC_\varepsilon|$ values are scaled by 1e-17.

Restrictions:

(a') = a with $\gamma^S = \gamma^U$

(a'') = a' with $\gamma^U = -\gamma^{K^e}$

(b): $\psi = 1$ plus "pure" Hicks;

(c): $\psi = 1$ with "identified" technical progress.

Table 3: Three Level Case, Constellation 2. in equation (6).

Specification:	$K^b, \psi, [U, \sigma, (K^e, \eta, S)]$		
Case:	(a)	(b)	(c)
ψ	0.866^{***}	1	1
σ	1.815^{***}	186.150	1.833^{***}
η	0.780^{***}	0.499^{***}	0.766^{***}
γ^{K^b}	0.012^{**}	–	–
γ^{K^e}	0.048^{***}	–	–
γ^{S}	0.020^{***}	–	–
γ^{U}	-0.015^{***}	–	–
γ^{H}	–	0.011^{***}	0.020^{***}
$\gamma^{K^e} - \gamma^S$	–	–	0.024^{**}
$\gamma^{U} - \gamma^S$	–	–	-0.035^{***}
	Restrictions and Diagnostics		
$\psi = 1$	[0.000]	–	–
$\sigma = 1$	[0.000]	[0.509]	[0.000]
$\eta = 1$	[0.002]	[0.000]	[0.001]
$\sigma = \eta$	[0.000]	[0.000]	[0.000]
$\gamma^i = \gamma^j \; \forall i, j$	[0.000]	–	–
$\gamma^{K^b} = \gamma^{K^e}$	[0.010]	–	–
$\gamma^{U} = \gamma^{S}$	[0.000]	–	–
	ADF tests		
FOC_{K^b}	[0.014]	[0.027]	[0.027]
FOC_{K^e}	[0.151]	[0.359]	[0.152]
FOC_{S}	[0.047]	[0.155]	[0.035]
FOC_{U}	[0.058]	[0.695]	[0.113]
CES	[0.006]	[0.141]	[0.006]
$\lvert VC_{\varepsilon} \rvert$	9.870	135.000	9.980

Note: See Notes to Table 2. Normalization parameters: $[\alpha_0, \beta_0, \pi_0] = [0.056, 0.635, 0.769]$.

Restrictions:

(b): $\psi = 1$ plus "pure" Hicks;

(c): $\psi = 1$ with "identified" technical progress.

Table 4: Three Level Case, Constellation 3. in equation (6).

Specification:	$K^b, \psi, [K^e, \sigma, (U, \eta, S)]$		
Case:	(a)	(b)	(c)
ψ	0.602^{***}	1	1
σ	0.647^{***}	1.158^{***}	0.933^{***}
η	2.338^{***}	9.984^{***}	2.568^{***}
γ^{K^b}	0.013^{***}	–	–
γ^{K^e}	-0.014^{***}	–	–
γ^S	0.029^{***}	–	–
γ^U	-0.004^{**}	–	–
γ^H	–	0.011^{***}	0.036^{***}
$\gamma^{K^e} - \gamma^S$	–	–	-0.111^{***}
$\gamma^U - \gamma^S$	–	–	-0.029^{***}
	Restrictions and Diagnostics		
$\psi = 1$	[0.000]	–	–
$\sigma = 1$	[0.000]	[0.000]	[0.009]
$\eta = 1$	[0.000]	[0.001]	[0.000]
$\sigma = \eta$	[0.000]	[0.001]	[0.000]
$\gamma^i = \gamma^j \ \forall i, j$	[0.000]	–	–
$\gamma^{K^b} = \gamma^{K^e}$	[0.000]	–	–
$\gamma^U = \gamma^S$	[0.000]	–	–
	ADF tests		
FOC_{K^b}	[0.004]	[0.027]	[0.027]
FOC_{K^e}	[0.002]	[0.003]	[0.001]
FOC_S	[0.023]	[0.393]	[0.003]
FOC_U	[0.007]	[0.974]	[0.059]
CES	[0.010]	[0.200]	[0.027]
$\lvert VC_\varepsilon \rvert$	1.160	4230.000	1.990

Note: See Notes to Table 2. Normalization parameters: $[\alpha_0, \beta_0, \pi_0] = [0.056, 0.856, 0.574]$.

Restrictions:

(b): $\psi = 1$ plus "pure" Hicks;

(c): $\psi = 1$ with "identified" technical progress.

Table 5: Contributions of Three Level Cases to the Skill Premium

Specification:	$K^b, \psi, [S, \sigma, (K^e, \eta, U)]$			$(K^b, \psi, U), \sigma, (K^e, \eta, S)$		$K^b, \psi, [K^e, \sigma, (U, \eta, S)]$	
Case:	(a'')	(a)	(c)	(a)	(c)	(a)	(c)
ω^S	-0.01390	-0.00836	-0.01267	-0.02446	-0.02449	-0.01442	-0.01313
ω^U	0.00209	0.00271	0.00190	0.00249	0.00246	0.00193	0.00176
$\boldsymbol{\omega^{RSS}}$	-0.0118	-0.00566	-0.01077	-0.02198	-0.02203	-0.01249	-0.01137
ω^{K^e}	-0.00261	-0.01664	-0.00230	0.00771	0.00801	–	–
ω^{Kb} †	–	–	–	–	–	–	–
$\boldsymbol{\omega^{KSC}}$	-0.00261	-0.01664	-0.00230	0.00771	0.00801	–	–
ω^{TC^S}	0.01545	0.01381	–	0.00557	–	0.01666	–
ω^{TC^U}	0.00606	-0.00529	–	0.00688	–	0.00223	–
$\omega^{TC^{K^e}}$	-0.00063	0.01905	–	0.00817	–	–	–
$\omega^{TC^{(K^e-S)}}$	–	–	0.00073	–	0.00432	–	0.00000
$\omega^{TC^{(U-S)}}$	–	–	0.01880	–	0.01613	–	0.01790
$\boldsymbol{\omega^{TC}}$	0.02081	0.02757	0.01953	0.02063	0.02045	0.01889	0.01790
$\widehat{g^{\omega}}$,†	0.00639	0.00527	0.00644	0.00637	0.00641	0.00640	0.00653

Note: † : *Data* : $g^{\omega} = 0.00632$. The (a) and (c) cases correspond to the earlier associated tables. Note, for compactness, we do not include the "pure" Hicks cases given that they represent the worst fit of all the exercises. † Unlike the two-level cases, there is no contribution to the skill premium from structures capital, given the separability.

Table 6: Growth Contributions: 3-Level CES cases

Specification:	$K^b,\psi,[S,\sigma,(K^e,\eta,U)]$			$(K^b,\psi,U),\sigma,(K^e,\eta,S)$		$K^b,\psi,[K^e,\sigma,(U,\eta,S)]$	
Case:	(a'')	(a)	(c)	(a)	(c)	(a)	(c)
y^S	0.01457	0.01461	0.01459	0.01451	0.01458	0.01455	0.01459
y^U	0.01770	0.00170	0.00178	0.00180	0.00181	0.00176	0.00178
$\boldsymbol{y^N}$	0.01634	0.01630	0.01637	0.01638	0.01638	0.01630	0.01638
y^{K^e}	0.00610	0.00625	0.00609	0.00603	0.00603	0.00625	0.00606
y^{K^b}	0.00139	0.00142	0.00135	0.00136	0.00135	0.00140	0.00135
$\boldsymbol{y^K}$	0.00749	0.00767	0.00744	0.00739	0.00738	0.00765	0.00741
y^{TC^S}	0.01206	0.00835	–	0.00914	–	0.01332	–
y^{TC^U}	−0.00396	0.00474	–	−0.00538	–	−0.00135	–
$y^{TC^{K^e}}$	0.00159	−0.00703	–	0.00652	–	−0.00189	–
$y^{TC^{K^b}}$	0.00089	0.00278	–	0.00069	–	0.00075	–
y^{TC^H}	–	–	0.02399	–	0.02007	–	0.03625
$y^{TC^{(U-S)}}$	–	–	−0.01142	–	−0.01249	–	−0.01024
$y^{TC^{(K^e-S)}}$	–	–	−0.00189	–	−0.00333	–	−0.01508
$\boldsymbol{y^{TFP}}$	0.01058	0.00903	0.01069	0.01097	0.01091	0.01083	0.01093
$\widehat{\boldsymbol{g^Y}},^\dagger$	0.03441	0.03300	0.03450	0.03474	0.03467	0.03478	0.03472

Note: † : *Data* : $g^Y = 0.03534$. See also note to table 5.

Table 7: Two Level Cases

Specification:	$\left(K^b,\zeta,S\right),\sigma,(K^e,\eta,U)$	$\left(K^b,\zeta,U\right),\sigma,(K^e,\eta,S)$	$\left(K^b,\zeta,K^e\right),\sigma,(U,\eta,S)$
	(a)	(b)	(c)
σ	2.951^{***}	2.505^{***}	0.527^{***}
η	1.697^{***}	1.317^{***}	2.078^{***}
ζ	1.234^{***}	7.327	0.816^{***}
γ^{K^b}	-0.060^{***}	-0.001	0.021^{***}
γ^{K^e}	0.012^{**}	-0.027^{***}	-0.016^{***}
γ^S	0.032^{***}	0.033^{***}	0.032^{***}
γ^U	-0.008^{***}	-0.003^{***}	-0.007^{***}
Restrictions and Diagnostics			
$\sigma=1$	[0.000]	[0.000]	[0.000]
$\eta=1$	[0.000]	[0.000]	[0.038]
$\sigma=\eta$	[0.000]	[0.000]	[0.001]
$\zeta=1$	[0.000]	[0.271]	[0.000]
$\gamma^i=\gamma^j\ \forall i,j$	[0.000]	[0.000]	[0.000]
$\gamma^{K^b}=\gamma^{K^e}$	[0.001]	[0.000]	[0.001]
$\gamma^U=\gamma^S$	[0.000]	[0.000]	[0.000]
ADF tests			
FOC_{K^b}	[0.465]	[0.213]	[0.105]
FOC_{K^e}	[0.016]	[0.088]	[0.01]
FOC_S	[0.016]	[0.013]	[0.044]
FOC_U	[0.055]	[0.146]	[0.002]
CES	[0.002]	[0.005]	[0.012]
$\|VC_\varepsilon\|$	2.460	2.760	2.340

Notes: Normalization parameters:

$$\left(K^b,\zeta,S\right),\sigma,(K^e,\eta,U)\quad:\quad[\alpha_0,\beta_0,\pi_0]=[0.480,0.294,0.888]$$
$$\left(K^b,\zeta,U\right),\sigma,(K^e,\eta,S)\quad:\quad[\alpha_0,\beta_0,\pi_0]=[0.600,0.231,0.854]$$
$$\left(K^b,\zeta,K^e\right),\sigma,(U,\eta,S)\quad:\quad[\alpha_0,\beta_0,\pi_0]=[0.191,0.709,0.429]$$

Table 8: Contributions of Two Level Cases to the Skill Premium

Specification:	$\left(K^b,\zeta,S\right),\sigma,(K^e,\eta,U)$	$\left(K^b,\zeta,U\right),\sigma,(K^e,\eta,S)$	$\left(K^b,\zeta,K^e\right),\sigma,(U,\eta,S)$
ω^S	-0.01336	-0.01640	-0.01623
ω^U	0.00182	0.00165	0.00217
$\boldsymbol{\omega^{RSS}}$	-0.01155	-0.01476	-0.01406
ω^{K^e}	-0.00323	0.00378	–
ω^{Kb}	0.00137	0.00090	–
$\boldsymbol{\omega^{KSC}}$	-0.00186	0.00468	–
ω^{TC^S}	0.01970	0.01726	0.01654
ω^{TC^U}	0.00442	0.00178	0.00349
$\omega^{TC^{K^e}}$	-0.00087	-0.00224	–
$\omega^{TC^{K^b}}$	-0.00324	-0.00003	–
$\boldsymbol{\omega^{TC}}$	0.02001	0.01677	0.02003
$\widehat{\boldsymbol{g^{\omega}}}$	0.00664	0.00660	0.00597

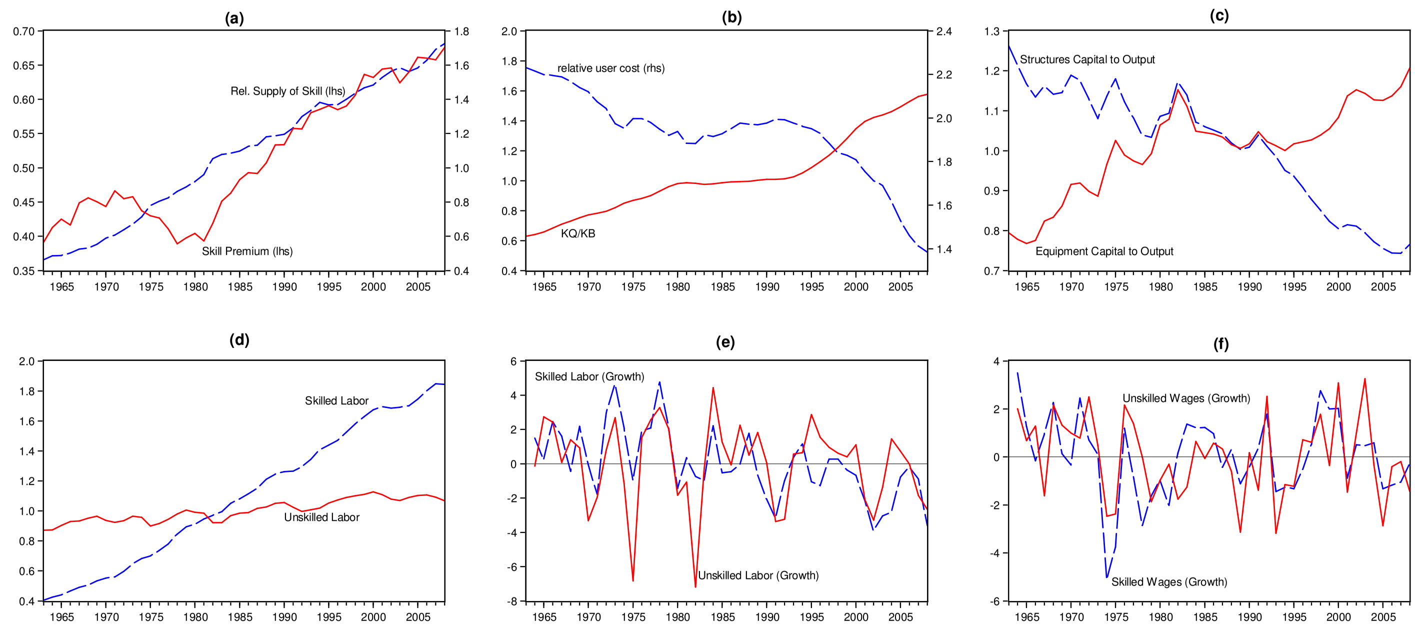
Note: See notes to Table 5.

Table 9: Growth Contributions: 2-Level CES cases

Specification:	$(K^b,\eta,S),\sigma,(K^e,\zeta,U)$	$(K^b,\eta,U),\sigma,(K^e,\zeta,S)$	$(K^b,\eta,K^e),\sigma,(U,\zeta,S)$
y^S	0.01457	0.01458	0.01444
y^U	0.00177	0.00178	0.00176
$\boldsymbol{y^N}$	0.01633	0.01636	0.01620
y^{K^e}	0.01633	0.00601	0.00637
y^{K^b}	0.00139	0.00138	0.00140
$\boldsymbol{y^K}$	0.00747	0.00747	0.00778
y^{TC^S}	0.01490	0.01534	0.01444
y^{TC^U}	-0.00259	-0.00096	-0.00234
$y^{TC^{K^e}}$	0.00160	-0.00363	-0.00224
$y^{TC^{K^b}}$	-0.00343	-0.00005	0.00123
$\boldsymbol{y^{TFP}}$	0.01048	0.01069	0.01109
$\widehat{\boldsymbol{g^Y}},^\dagger$	0.03429	0.03452	0.03506

Note: † : $Data : g^Y = 0.03534$.

Figure 1: The Skill Premium and Its Determinants



Notes: For ease of comparison, the growth in the skilled and unskilled labor input, as well as the growth in skilled and unskilled compensation are demeaned.

UNRAVELING THE SKILL PREMIUM:

ONLINE APPENDICES

Peter McAdam

European Central Bank and University of Surrey

Alpo Willman

University of Kent

A FIRST ORDER CONDITIONS ASSOCIATED WITH THE FOUR-FACTOR THREE-LEVEL CASE

that the representative firm faces an isoelastic demand curve, $Y_{it} = \left(\frac{P_{it}}{P_t}\right)^{-\varepsilon} Y_t$. Profit maximizing under the specified CES technology, (4), implies the following four first order conditions:[17]

$$\ln w_1 = \mathcal{C}_1 + \left(\frac{\psi-1}{\psi}\right)\gamma_1\tilde{t} + \frac{1}{\psi}\left[\ln\widetilde{Y} - \ln\widetilde{V}_1\right] \tag{A.1}$$

$$\begin{aligned}\ln w_2 = {} & \mathcal{C}_2 + \left(\frac{\sigma-1}{\sigma}\right)\gamma_2\tilde{t} + \frac{1}{\psi}\ln\widetilde{Y} - \frac{1}{\sigma}\ln\widetilde{V}_2 \\ & + \frac{(\psi-\sigma)}{\psi(\sigma-1)}\ln\left\{(1-\beta_0)\left(A_2\widetilde{V}_2\right)^{\frac{\sigma-1}{\sigma}} + \beta_0\left[(1-\pi_0)\left(A_3\widetilde{V}_3\right)^{\frac{\eta-1}{\eta}} + \pi_0\left(A_4\widetilde{V}_4\right)^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}\frac{\sigma-1}{\sigma}}\right\}\end{aligned} \tag{A.2}$$

$$\begin{aligned}\ln w_3 = {} & \mathcal{C}_3 + \left(\frac{\eta-1}{\eta}\right)\gamma_3\tilde{t} + \frac{1}{\psi}\ln\widetilde{Y} - \frac{1}{\eta}\ln\widetilde{V}_3 \\ & + \frac{(\psi-\sigma)}{\psi(\sigma-1)}\ln\left\{(1-\beta_0)\left(A_2\widetilde{V}_2\right)^{\frac{\sigma-1}{\sigma}} + \beta_0\left[(1-\pi_0)\left(A_3\widetilde{V}_3\right)^{\frac{\eta-1}{\eta}} + \pi_0\left(A_4\widetilde{V}_4\right)^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}\frac{\sigma-1}{\sigma}}\right\} \\ & + \frac{(\sigma-\eta)}{\sigma(\eta-1)}\ln\left\{(1-\pi_0)\left(A_3\widetilde{V}_3\right)^{\frac{\eta-1}{\eta}} + \pi_0\left(A_4\widetilde{V}_4\right)^{\frac{\eta-1}{\eta}}\right\}\end{aligned} \tag{A.3}$$

$$\begin{aligned}\ln w_4 = {} & \mathcal{C}_4 + \left(\frac{\eta-1}{\eta}\right)\gamma_4\tilde{t} + \frac{1}{\psi}\ln\widetilde{Y} - \frac{1}{\eta}\ln\widetilde{V}_4 \\ & + \frac{(\psi-\sigma)}{\psi(\sigma-1)}\ln\left\{(1-\beta_0)\left(A_2\widetilde{V}_2\right)^{\frac{\sigma-1}{\sigma}} + \beta_0\left[(1-\pi_0)\left(A_3\widetilde{V}_3\right)^{\frac{\eta-1}{\eta}} + \pi_0\left(A_4\widetilde{V}_4\right)^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}\frac{\sigma-1}{\sigma}}\right\} \\ & + \frac{(\sigma-\eta)}{\sigma(\eta-1)}\ln\left\{(1-\pi_0)\left(A_3\widetilde{V}_3\right)^{\frac{\eta-1}{\eta}} + \pi_0\left(A_4\widetilde{V}_4\right)^{\frac{\eta-1}{\eta}}\right\}\end{aligned} \tag{A.4}$$

where the individual constants are given by $\mathcal{C}_1 = \ln\left[\frac{\alpha_0}{(1+\mu)}\frac{Y_0}{V_{1,0}}\right]$, $\mathcal{C}_2 = \ln\left[\frac{(1-\alpha_0)(1-\beta_0)}{(1+\mu)}\frac{Y_0}{V_{2,0}}\right]$, $\mathcal{C}_3 = \ln\left[\frac{(1-\alpha_0)\beta_0(1-\pi_0)}{(1+\mu)}\frac{Y_0}{V_{3,0}}\right]$, $\mathcal{C}_4 = \ln\left[\frac{(1-\alpha_0)\beta_0\pi_0}{(1+\mu)}\frac{Y_0}{V_{4,0}}\right]$ and where $\mu = \varepsilon/(\varepsilon-1)$.

Denoting factor prices by w_i the distribution parameters α_0, β_0 and π_0 in equations (1)-(3) are defined by factor incomes at the normalization point,

$$\alpha_0 = (w_{1,0} \cdot V_{1,0}) \;/\; (w_{1,0} \cdot V_{1,0} + w_{2,0} \cdot V_{2,0} + w_{3,0} \cdot V_{3,0} + w_{4,0} \cdot V_{4,0}) \tag{A.5}$$

$$\beta_0 = (w_{3,0} \cdot V_{3,0} + w_{4,0} \cdot V_{4,0}) \;/\; (w_{2,0} \cdot V_{2,0} + w_{3,0} \cdot V_{3,0} + w_{4,0} \cdot V_{4,0}) \tag{A.6}$$

$$\pi_0 = (w_{4,0} \cdot V_{4,0}) \;/\; (w_{3,0} \cdot V_{3,0} + w_{4,0} \cdot V_{4,0}) \tag{A.7}$$

Equations (4)-(A.4) define a 5-equation system with manifest cross-equation parameter constraints. This encompasses the 3-equation system estimated by Krusell et al. (2000) who constrained the elasticity of substitution, ψ, between variable V_1 (structures capital) and the compound factor Z (capturing unskilled labor V_2, equipment capital V_3 and skilled labor V_4) to unity, i.e. Cobb Douglas.[18]

B FIRST ORDER CONDITIONS ASSOCIATED WITH $\psi = 1$ THREE-LEVEL CASE

The implied first order maximization conditions with respect to inputs corresponding (A.1)-(A.4) equations are:

$$\ln w_{1,t} = \mathcal{C}_1 + \ln\left[\frac{\alpha_0}{(1+\mu)}\frac{\widetilde{Y}_t}{\widetilde{V}_{1,t}}\right] \tag{B.1}$$

$$\begin{aligned} \ln w_{2,t} &= \mathcal{C}_2 + \left(\frac{\sigma-1}{\sigma}\right)\gamma_{24}t + \ln\left(\widetilde{Y}_t\right) - \frac{1}{\sigma}\ln\left(\widetilde{V}_{2,t}\right) \\ &- \ln\left\{\begin{array}{c} (1-\beta_0)\left(e^{\gamma_{24}t}\widetilde{V}_{2,t}\right)^{\frac{\sigma-1}{\sigma}} \\ +\beta_0\left[(1-\pi_0)\left(e^{\gamma_{34}t}\widetilde{V}_{3,t}\right)^{\frac{\eta-1}{\eta}} + \pi_0\left(\widetilde{V}_{4,t}\right)^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}\frac{\sigma-1}{\sigma}} \end{array}\right\} \end{aligned} \tag{B.2}$$

$$\begin{aligned} \ln w_{3,t} &= \mathcal{C}_3 + \left(\frac{\eta-1}{\eta}\right)\gamma_{34}t + \ln\left(\widetilde{Y}_t\right) - \frac{1}{\eta}\ln\left(\widetilde{V}_{3,t}\right) \\ &- \ln\left\{\begin{array}{c} (1-\beta_0)\left(e^{\gamma_{24}t}\widetilde{V}_{2,t}\right)^{\frac{\sigma-1}{\sigma}} \\ +\beta_0\left[(1-\pi_0)\left(e^{\gamma_{34}t}\widetilde{V}_{3,t}\right)^{\frac{\eta-1}{\eta}} + \pi_0\left(\widetilde{V}_{4,t}\right)^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}\frac{\sigma-1}{\sigma}} \end{array}\right\} \\ &+ \frac{(\sigma-\eta)}{\sigma(\eta-1)}\ln\left[(1-\pi_0)\left(e^{\gamma_{34}t}\widetilde{V}_{3,t}\right)^{\frac{\eta-1}{\eta}} + \pi_0\left(\widetilde{V}_{4,t}\right)^{\frac{\eta-1}{\eta}}\right] \end{aligned} \tag{B.3}$$

$$\begin{aligned} \ln w_{4,t} &= \mathcal{C}_4 + \ln\left(\widetilde{Y}_t\right) - \frac{1}{\eta}\ln\left(\widetilde{V}_{4,t}\right) \\ &- \ln\left\{\begin{array}{c} (1-\beta_0)\left(e^{\gamma_{24}t}\widetilde{V}_{2,t}\right)^{\frac{\sigma-1}{\sigma}} \\ +\beta_0\left[(1-\pi_0)\left(e^{\gamma_{34}t}\widetilde{V}_{3,t}\right)^{\frac{\eta-1}{\eta}} + \pi_0\left(\widetilde{V}_{4,t}\right)^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}\frac{\sigma-1}{\sigma}} \end{array}\right\} \\ &+ \frac{(\sigma-\eta)}{\sigma(\eta-1)}\ln\left[(1-\pi_0)\left(e^{\gamma_{34}t}\widetilde{V}_{3,t}\right)^{\frac{\eta-1}{\eta}} + \pi_0\left(\widetilde{V}_{4,t}\right)^{\frac{\eta-1}{\eta}}\right] \end{aligned} \tag{B.4}$$

where $\mathcal{C}_1 = \ln\frac{\alpha_0}{(1+\mu)}$, $\mathcal{C}_2 = \ln\left[\frac{\alpha_0(1-\beta_0)}{(1+\mu)}\frac{Y_0}{V_{2,0}}\right]$, $\mathcal{C}_3 = \ln\left[\frac{(1-\alpha_0)\beta_0(1-\pi_0)}{(1+\mu)}\frac{Y_0}{V_{3,0}}\right]$, $\mathcal{C}_4 = \ln\left[\frac{(1-\alpha_0)\beta_0\pi_0}{(1+\mu)}\frac{Y_0}{V_{4,0}}\right]$.

C FIRST ORDER CONDITIONS ASSOCIATED WITH THE TWO-LEVEL CASE

Isoelastic demand, and profit maximization, implies the first order conditions:

$$\begin{aligned}\ln w_{1,t} &= \mathcal{C}_1 + \frac{(\zeta-1)\,\gamma_1}{\zeta} t + \frac{1}{\sigma}\ln\left(\widetilde{Y}_t\right) - \frac{1}{\zeta}\ln\left(\widetilde{V}_{1,t}\right) \\ &+ \frac{\sigma-\zeta}{\sigma(\zeta-1)}\ln\left[(1-\beta_0)\left(e^{\gamma_1 t}\widetilde{V}_{1,t}\right)^{\frac{\zeta-1}{\zeta}} + \beta_0\left(e^{\gamma_2 t}\widetilde{V}_{2,t}\right)^{\frac{\zeta-1}{\zeta}}\right]\end{aligned} \tag{C.1}$$

$$\begin{aligned}\ln w_{2,t} &= \mathcal{C}_2 + \frac{(\zeta-1)\,\gamma_2}{\zeta} t + \frac{1}{\sigma}\ln\left(\widetilde{Y}_t\right) - \frac{1}{\zeta}\ln\left(\widetilde{V}_{2,t}\right) \\ &+ \frac{\sigma-\zeta}{\sigma(\zeta-1)}\ln\left[(1-\beta_0)\left(e^{\gamma_1 t}\widetilde{V}_{1,t}\right)^{\frac{\zeta-1}{\zeta}} + \beta_0\left(e^{\gamma_2 t}\widetilde{V}_{2,t}\right)^{\frac{\zeta-1}{\zeta}}\right]\end{aligned} \tag{C.2}$$

$$\begin{aligned}\ln w_{3,t} &= \mathcal{C}_3 + \frac{(\eta-1)\,\gamma_3}{\eta} t + \frac{1}{\sigma}\ln\left(\widetilde{Y}_t\right) - \frac{1}{\eta}\ln\left(\widetilde{V}_{3,t}\right) \\ &+ \frac{\sigma-\eta}{\sigma(\eta-1)}\ln\left[(1-\pi_0)\left(e^{\gamma_3 t}\widetilde{V}_{3,t}\right)^{\frac{\eta-1}{\eta}} + \pi_0\left(e^{\gamma_4 t}\widetilde{V}_{4,t}\right)^{\frac{\eta-1}{\eta}}\right]\end{aligned} \tag{C.3}$$

$$\begin{aligned}\ln w_{4,t} &= \mathcal{C}_4 + \frac{(\eta-1)\,\gamma_4}{\eta} t + \frac{1}{\sigma}\ln\left(\widetilde{Y}_t\right) - \frac{1}{\eta}\ln\left(\widetilde{V}_{4,t}\right) \\ &+ \frac{\sigma-\eta}{\sigma(\eta-1)}\ln\left[(1-\pi_0)\left(e^{\gamma_3 t}\widetilde{V}_{3,t}\right)^{\frac{\eta-1}{\eta}} + \pi_0\left(e^{\gamma_4 t}\widetilde{V}_{4,t}\right)^{\frac{\eta-1}{\eta}}\right]\end{aligned} \tag{C.4}$$

where the individual constants are given by $\mathcal{C}_1 = \ln\left[\frac{\alpha_0(1-\beta_0)}{(1+\mu)}\frac{Y_0}{V_{1,0}}\right]$, $\mathcal{C}_2 = \ln\left[\frac{\alpha_0\beta_0}{(1+\mu)}\frac{Y_0}{V_{2,0}}\right]$, $\mathcal{C}_3 = \ln\left[\frac{(1-\alpha_0)(1-\pi_0)}{(1+\mu)}\frac{Y_0}{V_{3,0}}\right]$, and $\mathcal{C}_4 = \ln\left[\frac{(1-\alpha_0)\pi_0}{(1+\mu)}\frac{Y_0}{V_{4,0}}\right]$.

Denoting factor prices by w_i $(i = 1, 2, 3)$ normalization implies that the distribution parameters α_0, β_0 and π_0 in (12)-(14) are defined by,

$$\alpha_0 = (w_{1,0}\cdot V_{1,0} + w_{2,0}\cdot V_{2,0})\,/\,(w_{1,0}\cdot V_{1,0} + w_{2,0}\cdot V_{2,0} + w_{3,0}\cdot V_{3,0} + w_{4,0}\cdot V_{4,0}) \tag{C.5}$$

$$\beta_0 = (w_{2,0}\cdot V_{2,0})\,/\,(w_{1,0}\cdot V_{1,0} + w_{2,0}\cdot V_{2,0}) \tag{C.6}$$

$$\pi_0 = (w_{4,0}\cdot V_{4,0})\,/\,(w_{3,0}\cdot V_{3,0} + w_{4,0}\cdot V_{4,0}) \tag{C.7}$$

Following Krusell et al. (2000) and others, we that $\widetilde{V}_1$ is structures capital, $\widetilde{V}_2$ unskilled labor, $\widetilde{V}_3$ equipment capital and $\widetilde{V}_4$ is skilled labor. Under this interpretation the two first equations (factor share equations) of the three equation system estimated by Krusell et al. (2000) are direct transformations of the first-order conditions (B.3)-(B.4). Their third equation (the rate of return equality condition), in turn, may be linked to the conditions (B.1)-(B.2). However, as they do not show its explicit derivation the possible correspondence remains ambiguous. As regards the underlying production function (4″) Krusell et al. (2000) left it outside their estimated 3-equation system.

D DECOMPOSITIONS: SKILL PREMIA AND OUTPUT

The individual growth contributions to output and the skill premium are derived as follows. We first have (using time subscripts for clarity),

$$\widehat{y}_t = ces\left(K_t^e, K_t^b, S_t, U_t \,;\, \widehat{\gamma}^{K^b}\tilde{t}, \widehat{\gamma}^{K^e}\tilde{t}, \widehat{\gamma}^{S}\tilde{t}, \widehat{\gamma}^{U}\tilde{t} \,;\, \widehat{\Sigma}\right) \tag{D.1}$$

$$\widehat{\omega}_t = f\left(K_t^e, K_t^b, S_t, U_t \,;\, \widehat{\gamma}^{K^b}\tilde{t}, \widehat{\gamma}^{K^e}\tilde{t}, \widehat{\gamma}^{S}\tilde{t}, \widehat{\gamma}^{U}\tilde{t} \,;\, \widehat{\Sigma}\right) \tag{D.2}$$

where $\widehat{y}$ and $\widehat{\omega}$ are, respectively, the estimated fits of the log of output and of the log skill premium conditional on the estimated technical change parameters, the substitution elasticity parameters, Σ, and on *ces* and f (the case-specific functional forms).

Then the growth contributions at time t, of, say, equipment capital to both output and the skill premium involve assuming that it remains at its previous level. Taking differences from estimated fits yield growth contributions of equipment capital to current period output and the skill premium from the previous period:

$$y_t^{K^e} = \widehat{y}_t - ces\left(K_{t-1}^e, K_t^b, S_t, U_t \,;\, \widehat{\gamma}^{K^b}\tilde{t}, \widehat{\gamma}^{K^e}\tilde{t}, \widehat{\gamma}^{S}\tilde{t}, \widehat{\gamma}^{U}\tilde{t} \,;\, \widehat{\Sigma}\right) \tag{D.3}$$

$$\omega_t^{K^e} = \widehat{\omega} - f\left(K_{t-1}^e, K_t^b, S_t, U_t \,;\, \widehat{\gamma}^{K^b}\tilde{t}, \widehat{\gamma}^{K^e}\tilde{t}, \widehat{\gamma}^{S}t, \widehat{\gamma}^{U}t \,;\, \widehat{\Sigma}\right) \tag{D.4}$$

Likewise, for the technical progress terms, the contribution of skill-augmenting technical change is found by assuming that there was no change between $t-1$ and t, i.e. replacing everywhere $\gamma^S \cdot \tilde{t}$ by

$\gamma^S \cdot (\tilde{t} - 1)$:

$$y_t^{TC^S} = \widehat{y}_t - ces\left(K_t^e, K_t^b, S_t, U_t\, ;\, \widehat{\gamma}^{K^b}\tilde{t}, \widehat{\gamma}^{K^e}\tilde{t}, \widehat{\gamma}^S\left(\tilde{t}-1\right), \widehat{\gamma}^U\tilde{t}\, ;\, \widehat{\Sigma}\right) \tag{D.5}$$

$$\omega_t^{TC^S} = \widehat{\omega}_t - f\left(K_t^e, K_t^b, S_t, U_t\, ;\, \widehat{\gamma}^{K^b}\tilde{t}, \widehat{\gamma}^{K^e}\tilde{t}, \widehat{\gamma}^S\left(\tilde{t}-1\right), \widehat{\gamma}^U\tilde{t}\, ;\, \widehat{\Sigma}\right) \tag{D.6}$$

Growth contributions of aggregated labor, capital, total factor productivity and total contribution of these factors are respectively obtained as (dropping time subscripts for convenience),

$$y^N = y^S + y^U \tag{D.7}$$

$$y^K = y^{K^e} + y^{K^b} \tag{D.8}$$

$$y^{TFP} = y^{TC^S} + y^{TC^U} + y^{TC^{K^e}} + y^{TC^{K^b}} \tag{D.9}$$

$$\widehat{g^Y} = y^N + y^K + y^{TFP} \tag{D.10}$$

Similarly the contributions of relative labor supply, capital skill complementarity, total technical change as well as their total contribution to the skill premium is obtained as,

$$\omega^{RSS} = \omega^S + \omega^U \tag{D.11}$$

$$\omega^{KSC} = \omega^{K^e} + \omega^{K^b} \tag{D.12}$$

$$\omega^{TC} = \omega^{TC^S} + \omega^{TC^U} + \omega^{TC^{K^e}} + \omega^{TC^{K^b}} \tag{D.13}$$

$$\widehat{g^\omega} = \omega^{RSS} + \omega^{KSC} + \omega^{TC} \tag{D.14}$$

In the context of three-level-CES production function the skill premium is independent from the development of structures capital stock and, therefore, $\omega_t^{K^b} = 0$

In the case of Four-Factor-Nested CD-CES production function total technical change (or TFP) contributions to output growth and to the growth of the skill premium are expressed in terms of common Hicks-neutral technical change component and the growth contributions of the deviations of unskilled labor and equipment capital augmenting technical changes from that of skilled labor augmenting technical change as follows,

$$y^{TFP} = y^{TC^H} + y^{TC^{(U-S)}} + y^{TC^{(K^e-S)}} \tag{D.15}$$

$$\omega^{TC} = \omega^{TC^H} + \omega^{TC^{(U-S)}} + \omega^{TC^{(K^e-S)}} \tag{D.16}$$

Tables 6 and 9 present average growth contributions of inputs and respective augmenting technical change components. The most striking is how similar growth contribution estimates each specification alternative gives. In all cases labor, capital and TFP have corresponded around 47%, 21-22% and 31-32%, respectively, of the growth in the sample period. Around 89% of labor contribution and around 82% of capital contribution except specification 1 of 2-level case where it is 92% – is related to the growth of skilled labor and the growth of equipment capital, respectively. Only the allocation TFP contributions varies across cases except the contribution of skilled labor augmenting technical change that uniformly exceeds somewhat that of the TFP implying somewhat negative net contributions from other augmenting technical components.

E Disaggregated Fit Measures of Estimations

Table E.1: Residual Standard Errors of The estimated Production Systems

	Table 2				
	(a)	(a')	(a'')	(b)	(c)
σ_{K_b}	0.140	0.136	0.083	0.094	0.100
σ_{K_e}	0.122	0.181	0.087	0.159	0.089
σ_S	0.043	0.041	0.027	0.136	0.022
σ_U	0.038	0.036	0.025	0.115	0.022
σ_Y	0.042	0.044	0.027	0.041	0.026
	Table 3				
	(a)			(b)	(c)
σ_{K_b}	0.097			0.097	0.100
σ_{K_e}	0.099			0.127	0.097
σ_S	0.036			0.038	0.034
σ_U	0.030			0.046	0.029
σ_Y	0.027			0.035	0.027
	Table 4				
	(a)			(b)	(c)
σ_{K_b}	0.081			0.094	0.100
σ_{K_e}	0.057			0.069	0.067
σ_S	0.035			0.058	0.022
σ_U	0.027			0.084	0.021
σ_Y	0.028			0.038	0.028
	Table 7				
	(a)			(b)	(c)
σ_{K_b}	0.128			0.131	0.091
σ_{K_e}	0.091			0.095	0.060
σ_S	0.029			0.028	0.048
σ_U	0.025			0.027	0.031
σ_Y	0.027			0.027	0.029

F HOW DOES THE SKILL PREMIUM WORK?

It is further interesting to note what our various systems work *out of sample*. To examine that, we extrapolate the exogenous variables in our five-equation system. We consider the following five scenarios:

1. $g^{\Lambda}_{T_+} = g^{\Lambda}$, "hist"

2. 1. except S : $g^{S}_{T_+} = g^{U}$

3. 1. except U : $g^{U}_{T_+} = g^{S}$

4. 1. except K^b : $g^{K^b}_{T_+} = g^{K^e}$

5. 1. except K^e : $g^{K^e}_{T_+} = g^{K^b}$

where $\Lambda = S, U, K^e, K^b$ and $T_+ \in [2009,\ 2050]$.

In scenario 1., we extrapolate all exogenous variables by their in-sample historical mean growth rates, g^{Λ}, forward to a symmetric future date, $T_+ = 2050$. The other scenarios do likewise but single out one particular growth pattern for special interest.[19]

Scenarios 2. and 3. eliminate the RSS effect, albeit in different ways. In 2. we set the out-of-sample growth rate of skilled labor, $g^{S}_{T_+}$, to the historical growth of unskilled labor, g^{U}. This implies a substantial reduction in the growth of skilled labor (recall Table 1). Scenario 3 reverses this, so that growth in unskilled labor becomes as high as skilled labor, which in turn implies a substantial increase in the growth of unskilled labor.

Scenario 4. raises the growth in structures capital to that of the rate of equipment capital. By contrast, in scenario 5. the growth in equipment capital falls to that of structures. Notice though that under projection scenarios 4. and 5., that the *ratio* of equipment to structures capital is the same in both cases (equal to its 2008 level) but that the individual levels in each scenario will differ:

$$\frac{g^{K^e}_{T_+} = g^{K^e}}{g^{K^b}_{T_+} = g^{Ke}} \equiv \frac{g^{K^e}_{T_+} = g^{K^b}}{g^{K^b}_{T_+} = g^{K^b}} : \begin{cases} 4. & \left(g^{K^b}_{T_+} = g^{K^e}\right) > \left(g^{K^b}_{T_+} = g^{K^b}\right) \\ \\ 5. & \left(g^{K^e}_{T_+} = g^{K^b}\right) < \left(g^{K^e}_{T_+} = g^{K^e}\right) \end{cases}$$

Figures F.1 and F.2 take the models in tables 3 and 4 as laboratories to examine extrapolated paths. These are the best performing models (the latter being the best). But also their choice is instructive since it illustrates systems with and without KSC, respectively. The figures plot, as before, the different projections for the log premium, ω, the relative log user cost, r, and the real GDP growth rate.

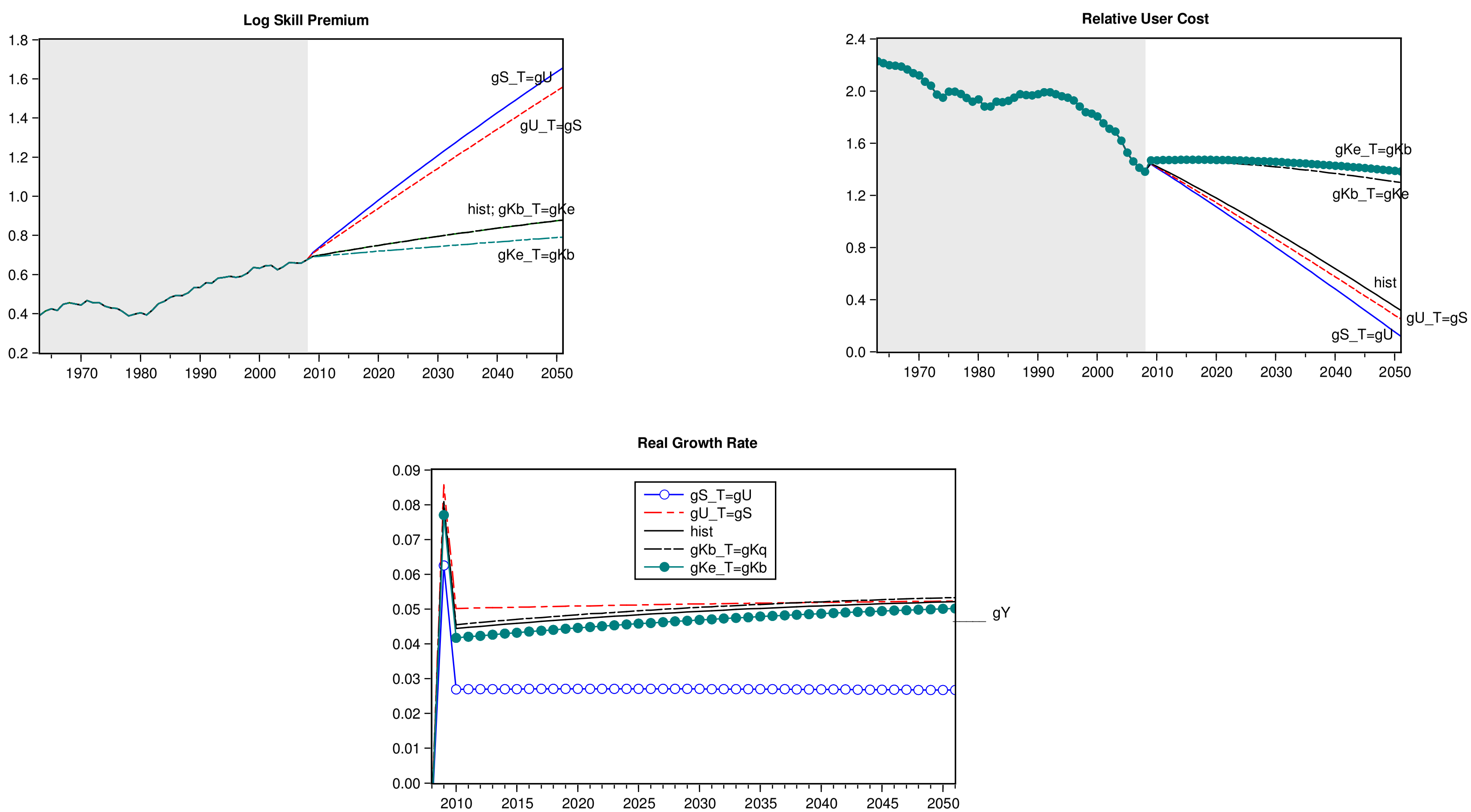
– Figures F.1 and F.2 here –

Projections based on *historical* trends are qualitatively similar across the two models. The log skill premium continues its upward trend (from around 0.68 in 2008 to 0.98 (model 4a), 0.88 (model 3a)). The relative user cost continues to decline on historical projections although with large differences across models (from around 1.4 to 0.8 (model 4), to 0.3 (model 3). The same is true for growth rates. Expanding factors will pull up growth since growth for factors is increasing, the differences in the growth trajectory reflect then the relative growth rates. Note since our framework is essentially static with constant positive factor growth rates, growth is thus acyclical and mechanically above its historical median.

Constellation 3 is the easiest to analyze since neither g^{K^e} nor g^{K^b} affect the premium. Moreover, the premium is always higher when $g^S = g^U$ or $g^U = g^S$ since the negative RSS component is removed entirely. Regarding user costs, in scenarios 4. and 5. the upward trend in K^e/K^b becomes fixed at its 2008 level. With capital more scarce relative to the other growing factors in the economy, user costs rise. In the case where the level of structures capital rises well above its historical growth rate (scenario 4.), r^{K^b} falls and widens the real relative user cost, $r^{K^e} - r^{K^b}$. The two labor scenarios lead to user cost trajectories similar to globally historical projections. On growth, where skilled labor grows only at the rate of unskilled labor, growth falls below that even of history. Likewise growth suffers when equipment capital is constrained to a lower growth rate. Otherwise, there are relatively small differences.

Let us now move to constellation 2. Notwithstanding qualitatively similar features, there are some large size differences with respect to model 3. Here, recalling equation (18), unless $g^S = g^U = 0$, the RSS will operate on the skill premium. Given this, we see high projections for the premium to be 0.1-0.2 higher in the labor series projections relative to the previous case. The path for relative user costs is, by comparison, much weaker. For the two capital growth expansions, the user cost term becomes essentially flat. Whilst for the historical and labor projections, the reduction in the user cost becomes more pronounced. In terms of growth, the outcomes are quite similar to before, except that now with KSC present, the $g^{K^e} = g^{K^b}$ has a less negative impact on growth relative to the historical trends case.

Figure F.1: Model Projections: Skill Premium, Relative Users Costs, Growth, Constellation 2: Model $K^b, \psi, [U, \sigma, (K^e, \eta, S)]$



Notes: The grey area denotes history. For real growth graphs, to maximize legibility, we started from the out-of-sample period; the g^Y line in the lowest panel is the median real growth rate, 0.039. Both this figure and figure F.2 are drawn on pairwise common axes for comparability.

Figure F.2: Model Projections: Skill Premium, Relative Users Costs, Growth, constellation 3: Model $K^b, \psi, [K^e, \sigma, (U, \eta, S)]$