Scale and Skill

Diego Comin* Matthew Dey† Martí Mestieri‡

March 20, 2025

Abstract

Using a unique dataset covering all US firms, we investigate a specific type of nonhomotheticities in production that links a firm's skill intensity to its scale. We document the presence of these non-homotheticities in all two-digit sectors except transportation and real estate. In 11 out of 17 non-homothetic sectors, skill intensity increases with the firm scale, while in the remaining 6 sectors, larger scale reduces skill intensity. We study how scale affects a firm's efficiency units per worker and contributes to variation in skill intensity. Scale is about as important as skill-biased technological change (SBTC) in explaining variation in skill intensity both in the cross section and the time series from 2000 to 2018. However, scale plays a more relevant role than SBTC in shaping the differential evolution in skill intensity between firms at the top and bottom quartiles of the firm growth distribution.

*Dartmouth College and NBER.
†BLS.
‡IAE-CSIC, CREI, BSE and CEPR.
We thank audiences from several institutions for useful comments and discussions. Mestieri acknowledges financial support from the Spanish Ministry of Science and Innovation Grant "Europa Excelencia 2024."

1 Introduction

Homothetic production functions are characterized by scale-independent relative factor shares, for given relative input prices. In non-homothetic formulations, instead, the relative factor intensity changes with the scale of operation, for given factor prices. Despite the wide-spread use of homothetic production functions, it is readily apparent that an expansion in scale can asymmetrically affect production factors. For example, an increase in the demand for specialized services may induce a firm to hire more skilled workers. In this case, scale leads to an increase in the skill intensity of the firm. Conversely, a firm may grow by adding more establishments while preserving the scale of the headquarters which may be more skill-intensive. In this case, scale leads to a reduction in the firm' skill intensity. The fact that these realistic illustrations are ruled out by the specifications posed in the majority of analyses suggests that the homotheticity of firm-level production is an understudied question.

In this paper, we explore a specific form of non-homotheticity in production, which introduces a scale effect on the mix of skilled and unskilled workers employed by a firm. To motivate our study, we consider two specific sectors: finance (NAICS code 521) and retail trade (NAICS codes 44-45). In each sector, we construct firm-level measures of the wage bill share of high-skilled workers, which we refer to as skill intensity, or simply skill. We also construct a proxy measure of scale by dividing the firm's wage bill by the average wage it pays to low-skilled workers. We then residualize the logarithms of these skill and scale measures by regressing them on the firm's relative wage of high- to low-skilled workers, four-digit industry dummies, and year dummies. Figures 1A and 1B show the bin-scatter plots of residual skill against residual scale for each sector (1A for finance and 1B for retail).

In a homothetic environment, there should be no relationship between these two variables. However, residual skill is strongly associated with residual scale for firms in both the finance and retail trade sectors. Interestingly, the associations differ markedly between the two sectors. While we observe a very strong positive association between skill and scale in finance, the association is strongly negative in retail.

These relationships, however, could result from omitted variable bias. For example, skilled-biased productivity differences across firms could be correlated with their scale, leading to a spurious association between scale and skill. It is unclear, though, why skilled-biased productivity would be positively associated with scale in finance and negatively in retail. A competing explana-

tion for the relationships between scale and skill in Figure 1 is that production is non-homothetic. In some sectors, as firms grow, production becomes more skill-intensive because growing firms require a larger proportion of skilled workers (e.g., more financial advisors and fewer secretaries). In other sectors, as firms grow, they require a larger share of unskilled workers (e.g., more storefront workers and relatively fewer store managers). Importantly, the variation in firm scale would be independent of the relative productivity of skilled versus unskilled workers in the firm.

To investigate the presence of non-homotheticities in production, we specify a flexible production structure that nests both homothetic and non-homothetic cases. The production function has two layers. The outer layer is a generalized Cobb-Douglas aggregator of capital and efficiency units of labor with potentially time-varying elasticities. The inner layer is an extension of the non-homothetic-CES aggregator in Hanoch (1975b) and Comin et al. (2021), which implicitly defines the number of efficiency units used in production as a function of the number of skilled and unskilled workers employed. Unlike Hanoch (1975b) and Comin et al. (2021), our aggregator allows for the possibility of time-varying taste parameters that capture skilled-biased technical change.

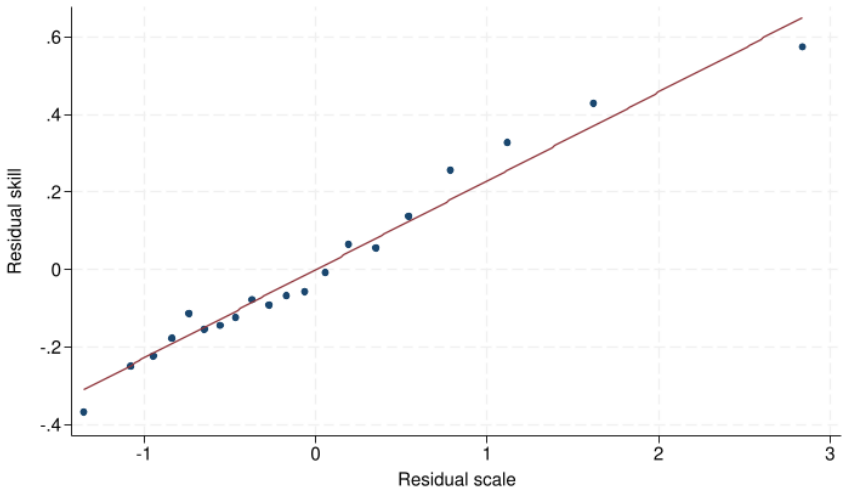
Using the first-order conditions, we derive a reduced-form expression for the (log) wage bill share of high-skilled relative to low-skilled workers. As in standard homothetic models, this depends on SBTC and on the relative wages of high- and low-skilled workers. Additionally, the non-homotheticities introduce a third driver of the relative wage bill share, which depends on the scale of the firm.

To identify the structural parameters of the production function for each sector, we construct six shift-share variables that instrument for the firm's relative wages and scale. The first instrument captures shifts in local labor supply induced by predicted migration flows from Latin America across locations. The second instrument reflects predicted demand shifts induced by changes in state-level median income in non-homothetic demand systems. The third and fourth instruments represent counterfactual wages for high- and low-skill workers paid by competitors in the same state.The final two instruments capture predicted changes in the population of individuals over 65 years old and female populations, using 20-year lagged demographic data at the county level. All these instruments are based on shifters that are common across firms in a given US location and that are apportioned based on firm-specific shares. The exogeneity of the instruments is therefore warranted by the exogenity of the shift components to the firm employment choices.

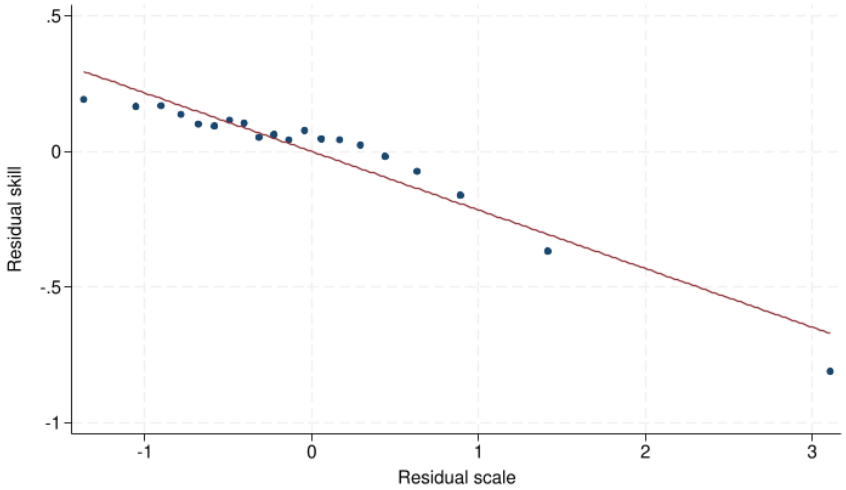
To estimate the model, we create a unique dataset that covers all firms in the US from 2000 to 2018. The dataset combines information on firm-level wage bills, number of employees, sec-

Figure 1: Scale and Skill. Partial Correlations for Two Industries

(a) Finance



(b) Retail



Notes: The plots report binned scatter plots of the partial correlations between scale and skill.

tor, and location from the Quarterly Census of Employment and Wages (QCEW) with detailed information on the number of employees and average salaries by occupation from the Occupational Employment and Wage Statistics (OEWS). We complement this data with information from the Census Bureau, the Bureau of Labor Statistics (BLS), and the Consumer Expenditure Survey (CEX) to construct the instruments.

Our model estimates reveal the systematic presence of non-homotheticities in production. In 18 out of the 19 2-digit sectors of the US economy, we reject the null hypothesis of homothetic

production, with transportation being the sole exception. Interestingly, there is significant heterogeneity across sectors both in terms of the magnitude and direction of non-homotheticities. In 11 of the 17 non-homothetic sectors, which we denote as skilled-elastic, an increase in the number of efficiency units of labor induces firms to become more skill-intensive. These sectors include mining, utilities, construction, wholesale trade, information, finance and insurance, professional services, education, healthcare, and arts, entertainment, and recreation. In the remaining six sectors, denoted as unskilled-elastic, an increase in the firm's scale leads to a reduction in skill intensity. Importantly, the estimates of skill-elasticity are robust to allowing SBTC to differ across firms in different quartiles/deciles of the firm-growth distribution. In addition to providing further reassurance about the exogeneity of the instruments, this result suggests that the causal effect of scale on skill identified in our baseline specification reflects the pressence of production non-homotheticities rather than endogenous responses of SBTC to changes in firm' scale.

Non-homotheticities in production have profound implications for a number of first order issues which include firm-level productivity, the dispersion in revenue-based measures of productivity across firms and the misallocation of factors of production, and the measurement of marginal costs and markups. We apply the estimated production functions to explore the implications of non-homotheticities for the number of efficiency units per worker and the skill intensity of firms.

The relationship between scale and the efficiency units of labor per worker depends on a parameter that also governs the relationship between marginal and average cost of efficiency units of labor. When the marginal cost is lower than the average cost, the firm faces increasing returns to scale and an increase in scale leads to an increase in the number of efficiency units per worker. The effect is the opposite when the marginal cost of efficiency units of labor is higher than the average cost. Our data is completely consistent with these model predictions.

We use our model, data, and estimates to evaluate the accuracy of the model predictions for firms' skill intensity, it's evolution and the drivers behind its variation. The model's explanatory power is remarkable. It accounts for 42% of the variance of (log) skill across firms. It precisely predicts the average increase in skill across the firms in each sector. Going beyond the average evolution of skill, We document that there is significant differences in the change in skill between fast- and slow-growing firms in a sector. Our model accounts for roughly half of this differential evolution of skill.

In homothetic environments, a firm' skill is determined by the level of skilled biased produc-

tivity and by the relative wages of skilled and unskilled workers. Non-homotheticities introduce a third driver: scale. We find that the two main drivers of the variation in skill across firms and of the average change in skill are skilled-biased technical change (SBTC) and scale, which are roughly equally important. However, scale is much more relevant than SBTC for the differential evolution of skill between fast- and slow-growing firms.

Related Literature Our analysis is related to several literatures. Our main contribution of documenting the presence of non-homotheticties in production is related to an extensive literature that has estimated production functions at the firm level. Studies in this area have targeted specific parameters and properties such as the returns to scale (e.g., Maksimovic and Phillips, 2001, 2002), the elasticity of substitution between capital and labor (Oberfield and Raval, 2021), the elasticity of substitution between high- and low-skilled workers (Katz and Murphy, 1992), or non-homotheticities that impact the capital intensity (Lashkari et al., 2024). The closest paper to ours is Trottner (2022) who studies the impact of expansions in sales driven by trade on the worker mix of German manufacturing establishments. In addition to some significant methodological differences we describe in section xx, our paper contributes to the literature in two dimensions. First, by covering all sectors of the economy, we can document the genericity of production non-homotheticities as well as the heterogeneity in sign and magnitude across sectors. Second, we consider that the relevant unit of analysis to study non-homotheticties is the firm rather than the establishment, as firms often grow by adding establishments and this process impacts the firm's worker mix. This may well may explain the different sign in the skill-elasticity we find for manufacturing.[1]

Non-homotheticities in production provide a theory of the efficiency units per worker which has implications for the measurement and estimation of productivity. This connects our paper to two large literatures. First, the dispersion in TFPR (Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009) which has interpreted variation across firms in TFPR as indication of markups and other distortions. Second, to the estimation of firm-level TFP, which typically relies on homothetic production structures (see De Loecker and Syverson, 2021 and the references therein)

More directly, our exploration of the implications of non-homotheticities for skill intensity is related to the literature on the drivers of inequality. This literature has differentiated between factors that affect the share of skill-intensive firms/sectors and factors that affect the skill intensity

[1]The estimation of the skill-elasticity is just one of the many contributions in Trottner (2022). For example, he develops a rich model of wage determination and studies the impact of trade expansions on both prices and quantities.

of firms within sectors. The first group includes trade (Autor et al., 2013), sectoral productivity differences (Barany and Siegel, 2019, 2020) and income growth (Comin et al., 2020). The second has been traditionally reduced to SBTC (e.g., Acemoglu and Autor, 2011). Our paper shows that in addition to SBTC, scale also affects the skill intensity of firms, and its omission provides a biased account of the drivers of inequality. WE show that scale is as important as SBTC in driving the average skill intensity of firms, and that differences is firm growth account for much of the variation in the evolution of skill intensity across the firms in any given sector.

The rest of the paper is structured as follows. Section 2 introduces a model of input choice when production is non-homothetic. Section 3 introduces the QCEW and OEWS datasets that we use to construct our data. Section 4 presents the estimation. Section 5 studies the model predictions for the drivers of efficiency units of labor per worker. Section 6 studies accuracy of the model predictions for skill across firms and its evolution as well as the role of various drivers of skill in the model. Section 7 concludes.

2 Model

2.1 Production

The output of firm f, which operates in sector s (Y_{fst}), is characterized by a generalized Cobb-Douglas function (1) that combines capital (K_{fst}) and efficiency units of labor (L_{fst}). Here, Z_{fst} represents the level of TFP in firm f, and β_{st} denotes the (time-varying) elasticity of output with respect to efficiency units of labor.

$$Y_{fst} = Z_{fst} L_{fst}^{\beta_{st}} K_{fst}^{1-\beta_{st}} \tag{1}$$

Efficiency units of labor are produced by combining workers of different skill levels or occupations according to the non-homothetic, constant elasticity of substitution (nh-CES) aggregator (2). Specifically, L_{fst} is an implicit function of the number of high- (H) and low-skilled (L) workers employed by firm f, denoted generically by L_{ofst}, where $o = \{H, L\}$. σ_s represents the elasticity of substitution between high- and low-skilled workers. $\widetilde{\varepsilon}_{os}$ governs the elasticity of demand of workers in occupation o with respect to the firm scale measured by the number of efficiency units,

$L_{f,s,t}$. When $\widetilde{\varepsilon}_{os} = 1 - \sigma_s \; \forall o$, (2) is the standard homothetic-CES.

$$1 = \left[\sum_o \left(\gamma_{ofst} L_{fst}^{\widetilde{\varepsilon}_{os}} \right)^{\frac{1}{\sigma_s}} \left(L_{ofst} \right)^{\frac{\sigma_s - 1}{\sigma_s}} \right]^{\frac{\sigma_s}{\sigma_s - 1}} \tag{2}$$

The nh-CES aggregator (2) allows for three parameter configurations. The first corresponds to $\sigma_s \in [0, 1)$ and $\widetilde{\varepsilon}_{os} > 0$. In this case, workers in different occupations of sector s are complements. The second configuration corresponds to $\sigma_s \in (1, \infty)$ and $\widetilde{\varepsilon}_{os} < 0$ in which workers in the different occupations of sector s are substitutes. The third corresponds to the case where the elasticity of substitution between occupations is unitary, $\sigma_s = 1$, and $\widetilde{\varepsilon}_{os} > 0$. In this case, (2) converges to[2]

$$\ln(L_{fst}) = \sum_o \frac{\gamma_{ofst} L_{fst}^{\widetilde{\varepsilon}_{os}} \ln(L_{fost})}{\sum_o \gamma_{ofst} L_{fst}^{\widetilde{\varepsilon}_{os}}}.$$

Without loss of generality, we can normalize one $\widetilde{\varepsilon}_{os}$ in each sector. Accordingly, when $\sigma_s \neq 1$, we set $\widetilde{\varepsilon}_{Ls} = 1 - \sigma_s$, and express $\widetilde{\varepsilon}_{Hs}$ as $\varepsilon_s^R \cdot (1 - \sigma_s)$, where $\varepsilon_s^R = \widetilde{\varepsilon}_{Hs} / \widetilde{\varepsilon}_{Ls}$. When $\sigma_s = 1$, we normalize $\widetilde{\varepsilon}_{Ls} = 1$ and $\widetilde{\varepsilon}_{Hs} = \varepsilon_s^R$. Importantly, ε_s^R and σ_s differ across sectors but are the same for the firms within each sector. We study the empirical support for this assumption below.

γ_{ofst} affects the intensity of occupation o in the production of efficiency units of labor and satisfies $\gamma_{ofst} > 0, \forall o$, and $\sum_o \gamma_{ofst} = 1$. Increases in γ_{ofst} reflect occupation-biased technical change caused, for example, by differential productivity growth of high- and low-skilled workers or by changes in the share of tasks conducted by high- and low-skilled workers (e.g., Zeira, 1998, Acemoglu and Autor (2011)).[3]

[2] See Appendix A.2

[3] γ_{ofst} captures occupation-biased technical change because, other things equal, it has a positive effect on the productivity of employment in occupation o, and a negative effect on the productivity of employment in occupation o'. Formally,

$$\frac{\partial^2 L_{fst}}{\partial L_{ofst} \partial \gamma_{ofst}} = \frac{1}{\sigma_s \gamma_{ofst}} \frac{\partial L_{fst}}{\partial L_{ofst}} \left(1 - \frac{\frac{\widetilde{\varepsilon}_{os}}{\sigma_s} \left(\gamma_{ofst} L_{fst}^{\widetilde{\varepsilon}_{os} - \sigma_s} \right)^{\frac{1}{\sigma_s}} \left(L_{ofst} \right)^{\frac{\sigma_s - 1}{\sigma_s}}}{\sum_o \frac{\widetilde{\varepsilon}_{os}}{\sigma_s} \left(\gamma_{ofst} L_{st}^{\widetilde{\varepsilon}_{os} - \sigma_s} \right)^{\frac{1}{\sigma_s}} \left(L_{ofst} \right)^{\frac{\sigma_s - 1}{\sigma_s}}}\right) > 0$$

where

$$\frac{\partial L_{fst}}{\partial L_{ofst}} = \frac{\frac{(1 - \sigma_s)}{\sigma_s} \left(\gamma_{ofst} L_{fst}^{\widetilde{\varepsilon}_{os}} \right)^{\frac{1}{\sigma_s}} \left(L_{ofst} \right)^{\frac{-1}{\sigma_s}}}{\sum_o \frac{\widetilde{\varepsilon}_{os}}{\sigma_s} \left(\gamma_{ofst} L_{fst}^{\widetilde{\varepsilon}_{os} - \sigma_s} \right)^{\frac{1}{\sigma_s}} \left(L_{ofst} \right)^{\frac{\sigma_s - 1}{\sigma_s}}} > 0$$

Similarly,

$$\frac{\partial^2 L_{fst}}{\partial L_{o'fst} \partial \gamma_{ofst}} = - \frac{\partial L_{fst}}{\partial L_{o'fst}} \frac{1}{\gamma_{o'fst}} \frac{\frac{\widetilde{\varepsilon}_{o's}}{\sigma_s} \left(\gamma_{o'fst} L_{fst}^{\widetilde{\varepsilon}_{o's} - \sigma_s} \right)^{\frac{1}{\sigma_s}} \left(L_{o'fst} \right)^{\frac{\sigma_s - 1}{\sigma_s}}}{\sum_o \frac{\widetilde{\varepsilon}_{os}}{\sigma_s} \left(\gamma_{ofst} L_{fst}^{\widetilde{\varepsilon}_{os} - \sigma_s} \right)^{\frac{1}{\sigma_s}} \left(L_{ofst} \right)^{\frac{\sigma_s - 1}{\sigma_s}}} < 0.$$

Differential increases in the productivity of an occupation can be caused either by different TFP processes or by

We start by modeling γ_{ofst} as the product of a firm random effect, ρ_{of}, and a sectoral time dummy, γ_{ost}, common to the firms in the sector. Later we consider formulations that allow for different time dummies across firms in different growth quartiles/deciles in a sector. In principle, the specific formulation of γ_{ofst} does not impact the estimation of the other structural parameters because we rely on exogenous variation induced by instrumental variables to identify them.

The nh-CES aggregator (2), is similar to that introduced by Hanoch (1975b) and used by Trottner (2022), with two important differences. First, the occupation-specific productivity parameters (γ_{ofst}) vary over time to reflect skilled-biased technical change. Second, the aggregator defines an index of efficiency units of labor instead of output. In section (4.1) we discuss the implications of this difference for the analysis.

2.2 Input demand

Let MC_{fst} be the marginal cost of production of firm f in sector s. Optimal input choice satisfies:

$$(1-\beta_{st})MC_{fst}Y_{fst} = R_t P_{Ct} K_{fst} \tag{3}$$

$$\beta_{st} MC_{fst} Y_{fst} = \widehat{W}_{fst} L_{fst} \tag{4}$$

where $\widehat{W}_{fst}$ is the marginal cost of an efficiency unit of labor in firm f. $\widehat{W}_{fst}$ is related to the average cost per efficiency unit of labor by

$$\widehat{W}_{fst} = \bar{\varepsilon}_{fst} \frac{WB_{fst}}{L_{fst}} \tag{5}$$

where $WB_{fst} = \sum_o W_{oft} L_{ofst}$ is the wage bill of firm f, and $\bar{\varepsilon}_{fst} = 1 + \alpha_{Hfst}(\varepsilon_s^R - 1)$ is the inverse of the scale elasticity (Hanoch, 1975b). Equation (5) shows that the scale elasticity introduces a wedge between the marginal and the average cost of efficiency units of labor. When $\bar{\varepsilon}_{fst} < 1$, the

capital-skill complementaries (e.g., Krusell et al., 2000). Our production structure captures capital-skilled complementarity by making γ_{ofst} a function of the amount of capital employed in production. That is $\gamma_{ofst} = \gamma_{o,s}(K_{fst})$ where $\partial\gamma_{H,s}(K)/\partial K > 0$, and $\partial\gamma_{L,s}(K)/\partial K < 0$. In this environment the FOC for capital differs from (3) below, as now the firm internalizes the impact that capital has on the number of efficiency units, given the number of skilled and unskilled workers employed. This effect does not affect our analysis which builds exclusively on the firm choices of high- and low-skilled workers.

In homothetic formulations of task-based production functions, $Y_{ft} = Z_{ft}\left[\gamma_{Ht}^{\frac{1}{\sigma}} L_{Hft}^{\frac{\sigma-1}{\sigma}} + \gamma_{Lt}^{\frac{1}{\sigma}} L_{Lft}^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$, where $\gamma_{Ht} = \frac{\int_0^{N_t^*}(\gamma(z)A_H)^{\sigma-1}dz}{\int_0^{N_t^*}(\gamma(z)A_H)^{\sigma-1}dz + \int_{N_t^*}^1(\gamma(z)A_L)^{\sigma-1}dz}$, $\gamma_L = (1-\gamma_H)$, where N_t^* denotes the threshold of tasks conducted by using skilled workers, and $Z_{ft} = \int_0^{N_t^*}(\gamma(z)A_H)^{\sigma-1}\,dz + \int_{N_t^*}^1(\gamma(z)A_L)^{\sigma-1}\,dz$.

marginal cost is lower than the average cost. In this case, the firm faces increasing returns to scale (IRS). When $\bar{\varepsilon}_{fst} > 1$, the marginal cost of efficiency units of labor is greater than the average cost inducing decreasing returns to scale (DRS). When $\bar{\varepsilon}_{fst} = 1$, marginal and average costs coincide and the firm faces constant returns to scale (CRS). Because $\alpha_{Hfst} \in (0,1)$, the returns to scale depend on whether ε_s^R is greater (DRS) or smaller (IRS) than 1. Note that, in our model, this parameter does not vary within a sector.

L_{fst} is implicitly defined by (2) and by the optimal mix of workers of each occupation in firm f, which is given by:

$$L_{ofst} = \gamma_{ost} \left(\frac{WB_{fst}}{W_{oft}} \right)^{\sigma_s} L_{fst}^{\widetilde{\varepsilon}_{os}} \tag{6}$$

The equivalent expression in the unit-elastic case:

$$L_{ofst} = \frac{WB_{fst}}{W_{oft}} \frac{\gamma_{ost} L_{fst}^{\widetilde{\varepsilon}_{os}}}{\sum_{o'} \gamma_{o'st} L_{fst}^{\widetilde{\varepsilon}_{o's}}} \tag{7}$$

2.3 Skill

Given (6), the share of occupation o in the wage bill of firm f is

$$\alpha_{ofst} \equiv \frac{W_{oft} L_{ofst}}{WB_{fst}} = \gamma_{ost} \left(\frac{WB_{fst}}{W_{oft}} \right)^{\sigma_s - 1} L_{fst}^{\widetilde{\varepsilon}_{os}} \tag{8}$$

Using (5) to replace WB_{fst}, we can express the share of high-skilled workers in the wage bill of firm f as

$$\alpha_{Hfst} = \overbrace{\gamma_{Hst}}^{SBTC} \overbrace{\left(\frac{\widehat{W}_{fst}/\bar{\varepsilon}_{fst}}{W_{Hft}} \right)^{\sigma_s - 1}}^{\text{Relative Wage effect}} \overbrace{L_{fst}^{(\varepsilon_s^R - 1)*(1-\sigma_s)}}^{\text{Non-homotheticities}} \tag{9}$$

This is our measure of skill intensity which we refer to as "skill". In the case where $\sigma_s = 1$, the firm skill is given by

$$\alpha_{Hfst} = \frac{\gamma_{Hfst} L_{fst}^{\varepsilon_s^R}}{\gamma_{Hfst} L_{fst}^{\varepsilon_s^R} + (1 - \gamma_{Hfst}) L_{fst}} \tag{10}$$

In our model, there are three potential drivers of the firm's skill (see equation (9)). As in the homothetic case, skill is affected by skilled-biased technical change and relative wages. Other things equal, the demand of high-skilled workers increases with SBTC. Additionally, when high-

and low-skilled workers are substitutes (i.e., $\sigma_s > 1$), higher skilled wages relative to the average cost of L_{fst} reduce the firm' skill, α_{Hfst}. In non-homothetic environments there is a third driver of skill in expression (9): scale. The number of efficiency units of labor affects the optimal mix of high- and low-skilled workers. In skill-elastic sectors ($(\varepsilon_s^R - 1) * (1 - \sigma_s) > 0$), other things equal, larger firms demand relatively more high-skilled workers. Conversely, in unskilled-elastic sectors ($(\varepsilon_s^R - 1) * (1 - \sigma_s) < 0$), other things equal, larger firms demand relatively more low-skilled workers. Therefore, whenever $\varepsilon_s^R \neq 1$ the firm' skill depends on the firm' scale and on the scale elasticity in the sector. In the case where $\sigma_s = 1$, the firm' skill is given by equation (10). As it is well known, in the unit-elastic case, relative wages do not affect skill intensity. However, scale may also impact skill. The skill-elasticity in this case is $\varepsilon_s^R - 1$.

2.4 Efficiency units per worker

The production aggregator (2) defines the number of efficiency units of labor in a firm. A natural question is how this differs from the number of workers employed in the firm. Next, we explore the relationship between the number of efficiency units L_{fst} and the total number of employees, $\overline{L}_{fst} \equiv \sum_o L_{fost}$. Using the production function aggregator, we express L_{fst} as

$$L_{fst} = \left[\sum_o \underbrace{\left(\gamma_{ost} L_{fst}^{\widetilde{\varepsilon}_{os} - (1-\sigma_s)} \right)^{\frac{1}{\sigma_s}}}_{\text{Weight}} \underbrace{\left(\frac{L_{ofst}}{\overline{L}_{fst}} \right)^{\frac{\sigma_s - 1}{\sigma_s}}}_{\text{Composition}} \right]^{\frac{\sigma_s}{\sigma_s - 1}} \overline{L}_{fst} \tag{11}$$

This expression shows that the number of efficiency units of labor per worker, $L_{fst}/\overline{L}_{fst}$ can be affected by factors that change the weights of skilled and unskilled workers in the aggregator or the composition of the firm's labor force. To gain further insights, we use the FOCs to express $\frac{L_{ofst}}{\overline{L}_{fst}}$ as

$$\frac{L_{ofst}}{\overline{L}_{fst}} = \frac{\gamma_{ost} \left(\frac{\widehat{W}_{fst}/\widetilde{\varepsilon}_{fst}^L}{W_{ot}} \right)^{\sigma_s^L} L_{fst}^{\widetilde{\varepsilon}_{os} + \sigma_s^L}}{\left(\frac{\widehat{W}_{fst}}{\widetilde{\varepsilon}_{fst}^L} \right)^{\sigma_s^L} \sum_o \frac{\gamma_{ost}}{W_{ot}^{\sigma_s^L}} L_{fst}^{\widetilde{\varepsilon}_{os} + \sigma_s^L}} = \frac{\frac{\gamma_{ost}}{W_{ot}^{\sigma_s^L}} L_{fst}^{\widetilde{\varepsilon}_{os} + \sigma_s^L}}{\sum_o \frac{\gamma_{ost}}{W_{ot}^{\sigma_s^L}} L_{fst}^{\widetilde{\varepsilon}_{os} + \sigma_s^L}} \tag{12}$$

Substituting in (11) we obtain

$$L_{fst} = \overline{L}_{fst} \left[\sum_o \left(\gamma_{ost} L_{fst}^{\widetilde{\varepsilon}_{os}-(1-\sigma_s)} \right)^{\frac{1}{\sigma_s}} \left(\frac{\frac{\gamma_{ost}}{W_{ot}^{\sigma_s^L}} L_{fst}^{\widetilde{\varepsilon}_{os}+\sigma_s^L}}{\sum_o \frac{\gamma_{ost}}{W_{ot}^{\sigma_s^L}} L_{fst}^{\widetilde{\varepsilon}_{os}+\sigma_s^L}} \right)^{\frac{\sigma_s-1}{\sigma_s}} \right]^{\frac{\sigma_s}{\sigma_s-1}}. \tag{13}$$

Imposing our normalization for the case of two occupations, (13) simplifies to:

$$\begin{aligned} L_{fst} = \overline{L}_{fst} \Bigg[& (1-\gamma_{Hst})^{\frac{1}{\sigma_s}} \left(\frac{1}{1+\frac{\gamma_{Hst}}{(1-\gamma_{Hst})} \left(\frac{W_{Lt}}{W_{Ht}} \right)^{\sigma_s^L} L_{fst}^{(\varepsilon_s^R-1)(1-\sigma_s)}} \right)^{\frac{\sigma_s-1}{\sigma_s}} \\ & + \left(\gamma_{Hst} L_{fst}^{(\varepsilon_s^R-1)(1-\sigma_s)} \right)^{\frac{1}{\sigma_s}} \left(\frac{\frac{\gamma_{Hst}}{(1-\gamma_{Hst})} \left(\frac{W_{Lt}}{W_{Ht}} \right)^{\sigma_s^L} L_{fst}^{(\varepsilon_s^R-1)(1-\sigma_s)}}{1+\frac{\gamma_{Hst}}{(1-\gamma_{Hst})} \left(\frac{W_{Lt}}{W_{Ht}} \right)^{\sigma_s^L} L_{fst}^{(\varepsilon_s^R-1)(1-\sigma_s)}} \right)^{\frac{\sigma_s-1}{\sigma_s}} \Bigg]^{\frac{\sigma_s}{\sigma_s-1}} \end{aligned} \tag{14}$$

Proposition 1 studies how scale and relative wages affect the gap between the number of efficiency units and the number of employees in a firm.

Proposition 1 *Suppose that $W_{Ht}/W_{Lt} > 1$. Other things equal, the number of efficiency units per worker in a firm, $L_{fst}/\overline{L}_{fst}$,*

a) increases with the relative wage of low-skilled workers, W_{Lt}/W_{Ht},

b) increases (decreases) with the scale, L_{fst}, if $\varepsilon_s^R < (>) 1$.

Proof: See Appendix.

The intuition of these results is as follows. Other things equal, relative wages affect $L_{fst}/\overline{L}_{fst}$ through the composition of employment but not through the weights of each occupation in the aggregator. An increase in W_{Lt}/W_{Ht} increases L_{Hfst}/L_{Lfst}. This leads to higher $L_{fst}/\overline{L}_{fst}$ if the marginal product of high-skilled workers is greater than low-skilled workers (i.e., $\frac{\partial L}{\partial L_H} > \frac{\partial L}{\partial L_L}$). But since in the optimum $\frac{\frac{\partial L}{\partial L_H}}{\frac{\partial L}{\partial L_L}} = \frac{W_{Ht}}{W_{Lt}}$, this is the case whenever $\frac{W_{Ht}}{W_{Lt}} > 1$. It is worth noting that part (a) of the proposition is irrespective of the value of ε_s^R. Therefore, it holds also in the homothetic case ($\varepsilon_s^R = 1$).

The intuition for part b is more convoluted because L_f affects both the ratio of skilled to unskilled workers, L_{Hfst}/L_{Lfst}, and the weight of skilled workers in the aggregator (14). As we show in the Appendix, the effect on the weight dominates and we focus on its intuition, next.

Other things equal, an increase in the weight of skilled workers leads to an increase in $L_{fst}/\overline{L}_{fst}$ if skilled and unskilled labor are substitutes (i.e. $\sigma_s > 1$), and to a reduction if they are complements. The skilled weight increases with L_f if $(\varepsilon_s^R - 1)(1 - \sigma_s) > 0$. Therefore, L_f increases the weight of skilled workers in (14) and this leads to an increase in $L_{fst}/\overline{L}_{fst}$ only when $\varepsilon_s^R < 1$. Similarly, an increase in scale leads to a reduction in the skilled workers weight if $(\varepsilon_s^R - 1)(1 - \sigma_s) < 0$. Other things equal, a reduction in the skilled workers weight leads to an increase in $L_{fst}/\overline{L}_{fst}$ if skilled and unskilled workers are complements (i.e. $\sigma_s < 1$). But for these two conditions to hold, $\varepsilon_s^R < 1$.[4] Part b of Proposition 1 implies that scale increases the number of efficiency units per worker only when there are increasing returns to scale (i.e. marginal cost of an efficiency unit of labor is smaller than the average cost).

3 Data

We rely on two Bureau of Labor Statistics' (BLS) data sources, the Occupational Employment and Wage Statistics (OEWS) program and the Quarterly Census of Employment and Wages (QCEW). The QCEW is built from the administrative state tax filings of the Unemployment Insurance System (supplemented with other information). It provides detailed industry (6-digit NAICS), ownership, geography (exact address), and overall employment and wage levels over time for every establishment in the United States that is covered by the Unemployment Insurance System. The OEWS sample is based on the QCEW frame. The program surveys roughly 200,000 establishments each May and November. OEWS respondents report employment counts by detailed occupation and wage intervals.

We exploit the detailed establishment characteristics available in both the OEWS and QCEW to either directly observe or impute occupational employment and wages for every establishment in QCEW. In brief, let

$$\Omega_t = R_t \cup N_t$$

where Ω_t represents the set of establishments in the QCEW at date t, R_t represents the set

[4]The intuition for the dominated effect (through the occupation mix) is as follows. An increase in the ratio L_{Hfst}/L_{Lfst} has a positive effect on $L_{fst}/\overline{L}_{fst}$ if the marginal product of skilled workers is larger than unskilled workers. However, unlike relative wages, the effect of scale on L_{Hfst}/L_{Lfst} depends on the scale elasticity. If $(\varepsilon_s^R - 1)(1 - \sigma_s) > 0$ an increase in scale increases the ratio of skilled to unskilled workers, and if $\frac{W_{Ht}}{W_{Lt}} > 1$ this leads to an increase in the number of efficiency units per worker.

of QCEW establishments that responded to the OEWS within the six closest OEWS panels, and N_t represents the set of QCEW establishments that did not respond to the OEWS over the same period, either because the unit was not sampled or it did not respond to the survey.[5] We directly observe OEWS outcomes for the units in R_t and we impute OEWS outcomes for units in N_t.

We employ a hierarchical imputation approach to find the single best donor based on how well OEWS respondents match non-responding QCEW frame units based on firm identify or the federal Employer Tax Identification Number (EIN), industry (six-digit NAICS), employment, and geography.

The first level of the hierarchy attempts to find a similarly sized OEWS respondent in the same firm and in the same six-digit industry.[6] Subsequent levels of the hierarchy attempt to find a similarly sized OEWS respondent in the same six-digit industry where we sequentially relax the level at which the frame unit and OEWS respondent match geographically from county to MSA to state and urban-rural status. At the lowest levels of the imputation hierarchy, we completely ignore geographic proximity, relax the level at which industries must match, and allow for larger employment differences between the frame unit and OEWS respondent.

For the private sector portion of the May 2017 QCEW frame, roughly 7% of units are OEWS respondents. These employers account for nearly 25% of total employment. In addition, nearly 16% of frame units accounting for 16% of total employment are imputed from similarly sized OEWS respondents in the same EIN and six-digit industry. In total, this means that slightly more than 40% of employment is assigned an occupation and wage based on data from a similarly sized OEWS respondent at the same firm in the same detailed industry.

The remaining data is generally imputed from a similarly sized OEWS respondent in the same six-digit industry and state. Nearly 73% of units and slightly more than 52% of employment is imputed using OEWS respondents at these levels of the hierarchy and, in fact, most of this data is imputed using matches within the same MSA. In addition, roughly 4% of units and 6% of employment is imputed from similarly sized OEWS respondents in the same six-digit industry, but outside of the state. Only about 0.4% of employment is imputed from an OEWS respondent

[5]The OEWS samples establishments at most once over six consecutive panels. Therefore, a full six panel OEWS surveys approximately 1.2 million unique establishments. For the purposes of constructing these population level datasets, we use OEWS respondents from the three panels prior, the concurrent panel, and two panels in the future. For example, for the May 2018 QCEW frame, we use OEWS data from November 2016, May and November 2017, May and November 2018, and May 2019.

[6]We define employment proximity as the absolute difference between the employment of the QCEW unit and the OEWS respondent divided by the average employment level of the two. We define similarly sized establishments when this measure is below 0.2. For example, if the frame unit has 100 workers, the OEWS respondent must be between 67 and 150 workers.

that does not match the frame unit at the six-digit NAICS level and only about 1% of employment is imputed from an OEWS respondent where employment levels differ by more than 20%.

Since imputed data for a QCEW frame unit may come from an OEWS respondent with a different employment level, we adjust occupational employment levels such that

$$L_{jet} = L_{jrt} \times \frac{L_{et}}{L_{rt}}$$

where L_{jet} represents employment in detailed occupation j at frame unit e in time t, L_{jrt} indicates occupation j employment at OEWS respondent r, and L_{et} and L_{rt} denote total employment at the frame unit and OEWS respondent, respectively. Essentially we keep the staffing pattern (i.e., occupational shares of employment) of the OEWS respondent fixed, but adjust employment levels to correspond to the size of the QCEW frame unit. Similarly, since the frame unit and OEWS respondent may be located in different geographic area or the OEWS unit may have responded at a past or future time period, we make geographic and time adjustments to the observed occupation-establishment wages so that the imputed wages of frame units are in local MSA and current dollars.

In addition to using the QCEW frame to define employment levels of establishments and subsequently firms, we also use the QCEW earnings data to define the total wage bill of firms. We define the total wage bill of firm-sector (roughly 2-digit NAICS with a couple exceptions) observations as the total annual earnings reported in the QCEW.

We map detailed occupations into our two highly aggregate occupations according to the typical education level required for entry into the occupation. Occupations that require a Bachelor's degree or higher are assigned to the high skill occupation and all other occupations are assigned to the low skill occupation.[7]

In our analysis, we restrict attention to firms with at least 20 employees with both high- and low-skilled workers. Additionally, we construct a balanced panel by requiring firms to appear in all the 7 periods we cover from 2000 to 2018. This restriction eliminates potential biases in the estimates by entry and exit (e.g., from different levels of skilled-biased productivity of firms that enter or exit the sample).

[7]See Education and training assignments by detailed occupation : U.S. Bureau of Labor Statistics (bls.gov) for the mapping between detailed occupation and required education level. In our mapping, we make ten exceptions where we map an occupation that does not require a Bachelor's degree into the high skill occupation. In all cases, these occupations require long-term on-the-job training. Almost all these occupations are in the Arts and Entertainment major occupation group. The list includes Athletes, Artists, Actors, Choreographers, Dancers, and Musicians. In total, these ten occupations account for 0.2% of total employment.

Table 1: Summary Statistics

	Sector	Obs.	$\frac{W_{H,f,t}}{W_{L,f,t}}$ 2000	$\frac{W_{H,f,t}}{W_{L,f,t}}$ 2018	$\frac{L_{H,f,t}}{L_{H,f,t}+L_{L,f,t}}$ 2000	$\frac{L_{H,f,t}}{L_{H,f,t}+L_{L,f,t}}$ 2018	$\alpha_{H,f,t}$ 2000	$\alpha_{H,f,t}$ 2018	Variance of Log: $\alpha_{H,f,t}$	Variance of Log: $\frac{WB_{f,s,t}}{W_{L,f,t}}$	Variance of Log: $\frac{W_{H,f,t}}{W_{L,f,t}}$
21	Mining	3822	1.96	2.30	0.22	0.26	0.36	0.45	0.71	1.64	0.11
22	Utilities	10668	1.54	1.50	0.19	0.27	0.27	0.35	0.31	1.42	0.05
23	Construction	71659	1.82	1.76	0.12	0.15	0.19	0.24	0.49	0.75	0.13
31	Manufacturing	161252	2.23	2.42	0.18	0.22	0.33	0.41	0.43	1.18	0.11
42	Wholesale Trade	73780	2.02	2.24	0.23	0.25	0.37	0.43	0.45	1.04	0.12
44	Retail Trade	69125	2.78	3.11	0.06	0.05	0.15	0.14	0.57	1.35	0.14
48	Transportation	20125	1.94	2.25	0.12	0.13	0.21	0.25	0.68	1.44	0.13
49	Warehousing	3885	1.62	1.85	0.08	0.08	0.13	0.14	0.52	1.94	0.10
51	Information (exc. Tech)	18396	1.56	1.84	0.32	0.36	0.42	0.51	0.35	1.54	0.10
518	Tech (see note below)	2604	2.04	1.84	0.67	0.75	0.80	0.85	0.18	1.88	0.07
52	Finance & Insurance	51100	2.08	2.10	0.37	0.50	0.55	0.68	0.27	1.46	0.07
53	Real Estate, Rental	11473	2.06	2.27	0.13	0.16	0.23	0.30	0.71	1.15	0.14
54	Prof., Sci., Tech. Serv.	96824	2.10	2.14	0.54	0.66	0.71	0.81	0.21	1.25	0.11
56	Admin., support serv.	25683	2.51	2.54	0.10	0.12	0.21	0.26	0.97	1.79	0.16
61	Educ. Serv. (exc. K-18)	19306	1.92	1.93	0.62	0.69	0.76	0.81	0.18	2.31	0.07
62	Health C., Social Assis.	170653	2.13	2.38	0.39	0.43	0.58	0.64	0.51	1.43	0.15
71	Arts, Entertain., Rec.	24003	2.41	2.55	0.15	0.15	0.30	0.31	0.72	0.89	0.16
72	Accomm., Food Serv.	7273	2.60	2.37	0.03	0.03	0.08	0.06	0.85	2.14	0.13
81	Other services	34846	2.09	2.38	0.25	0.28	0.41	0.48	0.79	0.88	0.13

Note: Mining is short for "Mining, Quarrying, and Oil and Gas Extraction "; Admin. & Support, Waste Mg't is short for "Administrative and Support and Waste Mananagment and Remediation Services. Columns 2-7 report the average value in each sector. Columns 8-10 report the variance of the log of the variable in each sector.

3.1 Descriptive statistics

Table 1 presents key summary statistics. Column 1 reports the number of firm. Columns 2-7 report the average across firms of each sector of the relative wages, the high-skilled share in employment and the wage bill.[8] Columns 8-10 report, for each sector, the variance across firms and periods of the (log) high-skill wage bill share, (log) wage bill over the wage of low skilled workers, and the log of the relative wages (high over low).[9]

There are a few relevant observations to note. First, in most sectors we observe an increase in the skill intensity measured by the wage bill share of high-skilled workers between 2000 and 2018. As is well known (e.g., Acemoglu and Autor, 2011), this increase is the result of both an increase in the relative wages of high- vs. low-skilled workers and in the share of high-skilled employment. Second, as documented by Barany and Siegel (2018) and Comin et al. (2020), there

[8] Averages are weighted by firm size measured by employment for relative wages and employment shares and by the wage bill for the wage bill share.

[9] When computing these variances, the observations are unweighted.

is significant heterogeneity across sectors in the increase in the high-skilled wage bill share. In sectors such as finance, management, information and insurance the increase has been greater than 10 percentage points, while in other sectors such as retail, accommodation or food services the share of the wage bill accrued by high-skilled workers has declined between 2000 and 2018. Third, there is considerable variation across firms in the log of skill intensity. This variation is larger than the cross-firm variance in the log of relative wages (high vs. low), but smaller than the cross-firm variance in the log of the wage bill over unskilled wage, which we use at this point as a proxy for firm size. These observations, together with the estimates of the structural parameters, help us assess the drivers of variation in skill across firms in section 6.1.

4 Estimation

We next discuss our empirical strategy to estimate the key structural parameters of the production function. The first step is to use the first-order conditions to derive an empirical specification. Then, we discuss the estimation strategy and the instruments we use to identify the model parameters. We then analyze the channel that drives the estimated causal effect of scale on skill. We conclude by studying the robustness of the estimates to the imputation process and their stability across the firm-size distribution.

4.1 Specification

Using the firm's optimal input choices, we have derived expression (9), which determines the firm's skill as a function of relative wages, scale, and model parameters. One difficulty of using this expression to estimate the model is that the firm scale, measured by the number of efficiency units of labor, is not observed in the data. To compute it directly using expression (2), we would need, in addition to information on the firm's employment in each occupation, the values of the parameters we intend to estimate γ_{ofst}, σ_s, and ε_s^R.

To resolve this, we follow Comin et al. (2021) and use the optimal intensity for low-skilled workers to express L_{fst} as a function of the variables we observe and model parameters. For the case where $\sigma_s \neq 1$, Equation (8) implies that L_{fst} can be expressed as

$$L_{fst} = \left(\frac{\alpha_{Lfst}}{\gamma_{Lfst}}\right)^{\frac{1}{1-\sigma_s^L}} \left(\frac{WB_{fst}}{W_{Lft}}\right). \tag{15}$$

Substituting in Equation (9), dividing both sides by α_{Lfst} and taking logarithms, we obtain an expression that relates the relative high- and low-skill shares to model parameters and observables only,

$$\ln\left(\frac{\alpha_{Hfst}}{\alpha_{Lfst}}\right) = \left[\ln(\gamma_{Hfst}) - \varepsilon_s^R \ln(\gamma_{Lfst})\right] + \left(\varepsilon_s^R - 1\right)\ln(\alpha_{Lfst}) + (1-\sigma_s^L)\cdot\left(\varepsilon_s^R - 1\right)\ln\left(\frac{WB_{fst}}{W_{Lft}}\right) + (1-\sigma_s^L)\cdot\ln\left(\frac{W_{Hft}}{W_{Lft}}\right). \tag{16}$$

Applying the same procedure when $\sigma_s = 1$, the log relative intensity is given by

$$\ln\left(\frac{\alpha_{Hfst}}{\alpha_{Lfst}}\right) = \ln\left(\frac{\gamma_{Hfst}}{\gamma_{Lfst}}\right) + (\varepsilon_s^R - 1)\ln\left(\frac{WB_{fst}}{\overline{W}_{fst}}\right) \tag{17}$$

where $\overline{W}_{fst} = \prod_o \left(\frac{W_{ofst}}{\alpha_{ofst}}\right)^{\alpha_{ofst}}$ is the average wage index in the unit elastic case.[10]

Before presenting the procedure to estimate equations (16) and (17), we compare our control approach to deal with the non-observability of L_{fst} with other approaches used in the literature. An alternative proposed by Hanoch (1975a) used by Trottner (2022) is that the non-homothetic index defined by the aggregator is a variable observed in the data, such as real output. This approach presents some challenges. First, real output is rarely observed at the firm level. Second, even if it is observed it is not very useful as conventional price index used to deflate nominal output differ from the non-homothetic price index required to define the model consistent measure of the output index (see Comin et al., 2021.) Additionally, trying to directly construct output from the aggregator is more challenging than computing efficiency units of labor, as it requires properly specifying the production function including all the inputs and measuring them. In contrast, computing the number of efficiency units of labor using (2) is straightforward, given all parameter values and the number of workers employed in each occupation.

4.2 Estimation procedure

Equations (16) and (17) provide the theoretical foundation for our estimating equations. Our empirical model allows for common time shocks across observation units in the form of time$\times$ sector fixed effects δ_{st}, firm random effects δ_f, and additively separable measurement error ν_{fst} to yield,

[10] Note also that when $\sigma_s = 1$, $WB_{fst} = L_{fst} * \overline{W}_{fst}$.

for the case $\sigma \neq 1$,

$$\ln\left(\frac{\alpha_{Hfst}}{\alpha_{Lfst}}\right) = \phi_{1s}\ln(\alpha_{Lfst}) + \phi_{2s}\ln\left(\frac{WB_{fst}}{W_{Lft}}\right) + \phi_{3s}\ln\left(\frac{W_{Hft}}{W_{Lft}}\right) + \delta_{st} + \delta_f + \nu_{fst}. \tag{18}$$

Our theory imposes the non-linear constraint that $\phi_{1s} = \frac{\phi_{2s}}{\phi_{3s}}$ and further implies that the terms capturing sectoral SBTC are subsumed in $\delta_{st} + \delta_f$. Given the non-linear nature of the constraint and the fact that $\ln(\alpha_{Lfst})$ appears on both sides of the estimating equation, introducing potential bias in our estimates, we proceed by estimating our model by two-stage least squares using the Iterated Linear Lease Squares (ILLE) procedure proposed in Blundell and Robin (1999).[11]

This estimation strategy allows us to exploit the conditional linearity of the model as follows. We start with a guess for $\phi_{1s}^{(0)}$ for each sector s and compute

$$\widetilde{\ln\left(\frac{\alpha_{Hfst}}{\alpha_{Lfst}}\right)}^{(0)} \equiv \ln\left(\frac{\alpha_{Hfst}}{\alpha_{Lfst}}\right) - \phi_{1s}^{(0)}\ln(\alpha_{Lfst}). \tag{19}$$

We then estimate via 2SLS the linear regression

$$\widetilde{\ln\left(\frac{\alpha_{Hfst}}{\alpha_{Lfst}}\right)}^{(0)} = \phi_{2s}\ln\left(\frac{WB_{fst}}{W_{Lft}}\right) + \phi_{3s}\ln\left(\frac{W_{Hft}}{W_{Lft}}\right) + \delta_{st} + \delta_f + \nu_{fst}, \tag{20}$$

to obtain a new estimate of $\phi_{1s}^{(1)}$ using that our structural model implies that $\frac{\phi_{2s}}{\phi_{3s}} = \phi_{1s}$, where we instrument for both the wage bills and relative wages (as we discuss further below). We use the new estimate $\phi_2^{(1)}$ to construct $\widetilde{\ln\left(\frac{\alpha_{Hfst}}{\alpha_{Lfst}}\right)}^{(1)}$ and estimate again Equation (20) via 2SLS. We iterate this procedure until the ϕ_{1s} converges. In practice, this convergence is very fast and requires less than five iterations.[12] Blundell and Robin (1999) show that this procedure generates consistent estimates of Equation (20).

The empirical specification in the unit-elastic case is

$$\ln\left(\frac{\alpha_{Hfst}}{\alpha_{Lfst}}\right) = \phi_{1s}\ln\left(\frac{WB_{fst}}{\overline{W}_{fst}}\right) + \delta_{st} + \delta_f + \nu_{fst}. \tag{21}$$

[11] In the online appendix of Comin et al. (2021), they discuss the use of ILLE as an alternative method for estimating the nonhomothetic CES preferences, and show that it performs equally as their baseline method in Montecarlo simulations. Blundell and Robin (1999) show the equivalence of ILLE with three-stage non-linear least squares and that it can be used in the presence of endogenous regressors.

[12] We define convergence if the difference between two iterations, n and $n+1$ satisfy that $|(\phi_{1s}^{(n+1)} - \phi_{1s}^{(n)})|/\phi_{1s}^{(n)}$ is less than 10^{-5}.

In this case, we can use standard 2SLS to estimate the model parameters. In particular, $\varepsilon_s^R = 1 + \hat{\phi}_{1s}$.

4.3 Instruments

One challenge in identifying the structural parameters by the described procedure is the possibility that firms experience different rates of skill-biased technical change that are correlated with firm scale or relative wages. This would bias the estimates of ϕ_{2s} and ϕ_{3s}. To ensure that the identification of these parameters in our estimating equations (20 and 21) is based on exogenous variation in firm scale and relative wages, we instrument firm scale and relative wages with six instruments established in the literature.

Our instruments fall under the broad class of shift-share instruments. The validity of these instruments hinges on the exogeneity of the shifts or shares, as discussed in Adao et al. (2019) and Borusyak et al. (2022). We construct the instruments by finding exogenous shifters common across U.S. locations and apportioning them based on firm-specific shares. Thus, the exogeneity of our resulting instruments is warranted by the exogeneity of the shift components to the firm employment choices. We next discuss the construction of each of these instruments and their validity.

Local Labor Supply Shocks The first instrument uses shifts in local labor supply and it has been widely used in the migration literature, e.g., Peri and Yasenov (2019) and Jaeger et al. (2018) among others. In particular, we construct Bartik-style predicted migration flows from Latin America across firm locations as a labor supply shock. The predicted number of Latin American workers available to the firm is constructed as

$$Z_{1fst} = \sum_{g \in G} \frac{\overline{L}_{gfst}}{\overline{L}_{fst}} \widehat{N}_{gt}, \tag{22}$$

where $\overline{L}_{gfst}$ is the employment level of firm f in sector s in year t located in state g, $\overline{L}_{fst}$ represents total employment of the firm-sector-year, $\widehat{N}_{gt}$ is the predicted level of Latin American workers in state-year gt and G denotes a set containing all US states. The predicted level of Latin American workers leverages the fact that migrants tend to go where other migrants live and it is based on the product of two components estimated from the Current Population Survey (CPS). The first component is the share of statewide employment between 2000 and 2002 performed by workers

from Latin America. The second component is the U.S. working-age population born in Latin American countries (N_t).[13] This instrument reflects a standard shock to the local labor supply, which is arguably exogenous to existing firms. We deem reasonable to assume that Latin American workers' choice of their state of residence is influenced by the initial share of Latin American migrants in the state.

Local Demand Shocks The second instrument captures predicted sectoral demand shifts due to nonhomotheticities in demand induced by changes in the state-level median income . We proceed in three steps to construct this instrument. First, We construct a measure, $\widehat{C}^{VA}_{gst}$, of the sectoral value added in each sector-state pair induced by changes in state-level average income and nonhomotheticities in demand, while holding relative prices constant. To construct such variable we use household-level estimates of the income elasticity of sectoral demand from the Consumer Expenditure data (CEX),[14] data on state-level median income, and a version of the nonhomothetic demand system in Comin et al. (2021) estimated to eight sectors. Relative prices are kept at the levels in the initial period. Second, we multiply $\widehat{C}^{VA}_{gst}$ by the sector's wage bill share in value added in 1980 in the US. This reflects the wage bill induced by the evolution of income through changes in sectoral demand (WB_{s0}/VA_{s0}). Third, we construct the demand instrument for firm f that operates in a sub-sector of s as the weighted average of $\frac{WB_{s0}}{VA_{s0}}\widehat{C}^{VA}_{gst}$ across all states where it operates, with weights given by the firm's employment share in each state. Formally,

$$Z_{2fst} = \sum_{g} \frac{\overline{L}_{gfst}}{\overline{L}_{fst}} \frac{WB_{s0}}{VA_{s0}} \widehat{C}^{VA}_{gst}. \tag{23}$$

Since changes in aggregate state consumption are exogenous to the firm (conditional on the time fixed effect), this instrument satisfies the exclusion restriction and it provides a demand shifter that is exogenous to the firm.

Counterfactual Wages The third and fourth instruments capture counterfactual wages for high- and low-skill workers paid by competitors in the same state. We use the leave-out average wages paid by firms in the same state and detailed occupation as a shift, and the employment share of the

[13]The CPS sample is too small in some states to rely on a single year of data. Therefore, we pool the monthly data over three years to get a more accurate picture of the employment share accounted for by Latin American immigrants in the base year. We include the countries of Mexico, the Caribbean, Central America, and South America in our definition of Latin America.

[14]Specifically, we estimate these elasticities for 8 broad sectors. See Comin et al. (2020) for details on the methodology and estimates.

detailed occupation and state in the firm's total employment in the initial sample year as a share. Therefore, these instruments reflect the average wage a firm would expect to pay for occupation H or L based on the prevailing detailed occupation wages in other firms in the sector and state and the firm's own occupation mix. Formally, we construct our instruments as

$$Z_{3fst} = \sum_{g} \sum_{j:o(j)=H} \frac{L_{jgfs0}}{L_{Hfs0}} \widetilde{w}_{jgfst}, \tag{24}$$

and

$$Z_{4fst} = \sum_{g} \sum_{j:o(j)=L} \frac{L_{jgfs0}}{L_{Lfs0}} \widetilde{w}_{jgfst}, \tag{25}$$

where $o(j)$ is the mapping from detailed occupation j into an aggregate occupation, L_{jgfst} represents employment by detailed occupation, state, firm, sector, and year, L_{Hfst} is the firm-sector-year employment in aggregate occupation H, L_{Lfst} is firm-sector-year employment in aggregate occupation L, and $\widetilde{w}_{jgfst}$ is the average wage paid to workers by detailed occupation, state, sector, and year, excluding firm f.

The shift component of these two instruments are the state-level detailed occupation wages (leaving out the own firm). Therefore, to the extent that no firm has a sizable effect on the equilibrium wage of a detailed-level occupation, this constitutes a plausible exogenous shifter from the firm's point of view since it captures aggregate changes in wages that are exogenous to the firm.

Demographic Composition The next two instruments capture demand changes that are driven by secular trend in the economy, and are related to changing demographic composition. Our demographic composition instruments leverage long-term changes in the age distribution and gender. In particular, the shift components are based on estimates of the population aged 65 and older and the female population at the county level using the 20-year lagged population over 45 years old and the female population, while the share uses firm exposure to different US counties. More concretely,

$$Z_{5,fst} = \sum_{c \in C} \left(\hat{P}^{(45+)}_{ct-20} \cdot \frac{L_{fsct}}{L_{fst}} \right)$$

$$Z_{6,fst} = \sum_{c \in C} \left(\hat{P}^{F}_{ct-20} \cdot \frac{L_{fsct}}{L_{fst}} \right)$$

where $\hat{P}^{(45+)}_{ct-20}$ and $\hat{P}^{F}_{ct-20}$ are the population aged 45 and older and the female population in county c twenty years ago, $\frac{L_{fsct}}{L_{fst}}$ denotes the employment share of firm f in sector s at time t in county c,

and C denotes the set of US counties.

These are standard shift-share population-composition instruments that capture predetermined trends in demand that are exogenous to any one firm (e.g., Jaravel, 2019). To satisfy the exclusion restriction, we need that these changes in demographic county characteristics are orthogonal to any specific firm. We deem plausible that there is no one single firm that drives the demographic changes in a county, and therefore the exclusion restriction is satisfied.

Strength of instruments Table A1 presents the J-statistics for the null that the instruments are not relevant in predicting each of the instrumented variables, $\frac{W_{Hft}}{W_{Lft}}$ and $\frac{WB_{ft}}{W_{Lft}}$ when $\sigma_s \neq 1$ and $\frac{WB_{ft}}{\overline{W}_{ft}}$ when $\sigma_s = 1$, in each sector. In all cases, the J-statistics are high and we can reject the null at any relevant level of significance.

4.4 Estimates

Table 2 presents the estimates of the structural parameters for each sector. Columns 1-4 show the point estimates of σ_s and ε^R with their standard errors.[15] Column 5 reports the estimated growth rate of $\gamma_{H,t}$ from 2000 to 2018, (i.e. the estimate of SBTC).[16] Column 6 shows the skill elasticity, calculated as $(\varepsilon^R - 1) * (1 - \sigma_s)$, when $\sigma_s \neq 1$, and as $\varepsilon^R - 1$ when σ_s is unitary. A positive value of the skill elasticity implies that, other things equal, the wage bill share of high-skilled workers, α_{Hfst}, increases with the number of efficiency units of labor, L_{fst}. In this case, we say that the sector is skilled-elastic. A negative value of the scale elasticity implies that, other things equal, α_{Hfst} decreases with L_{fst}. In this case, we say that the sector is unskilled-elastic.

The model estimates reveal two important findings. First, with the exceptions of transportation and real estate, ε_s^R is statistically different from one in all sectors. Thus, in all firms except those in transportation and real estate, the occupational mix of employment is affected by the firm' scale. Second, there is wide variation across sectors in the direction and magnitude of scale elasticities. To explore this, Figure 2 plots the estimated values of ε_s^R (horizontal axis) and σ_s (vertical axis) for each sector. The vertical line $\varepsilon_s^R = 1$ divides sectors between those with increasing returns to scale (marked as IRS) where $\varepsilon_s^R < 1$, and sectors with decreasing returns to scale (DRS) where $\varepsilon_s^R > 1$. Additionally, the loci $\sigma_s = 1$ and $\varepsilon_s^R = 1$ define four quadrants. We refer to them using the numbers

[15] We impose a lower bound for σ_s of 0.15. When the algorithm converges to a value smaller than this lower bound, we calibrate $\sigma_s = .15$ and proceed with the estimation. Similarly, when the estimated value tends to 1, we calibrate $\sigma_s = 1$ and proceed to estimate the model using the expressions for the unit-elasticity case.

[16] See Table A1 in the appendix for the estimates of $\gamma_{H,t}$ in each year. In most sectors, its evolution is monotonic.

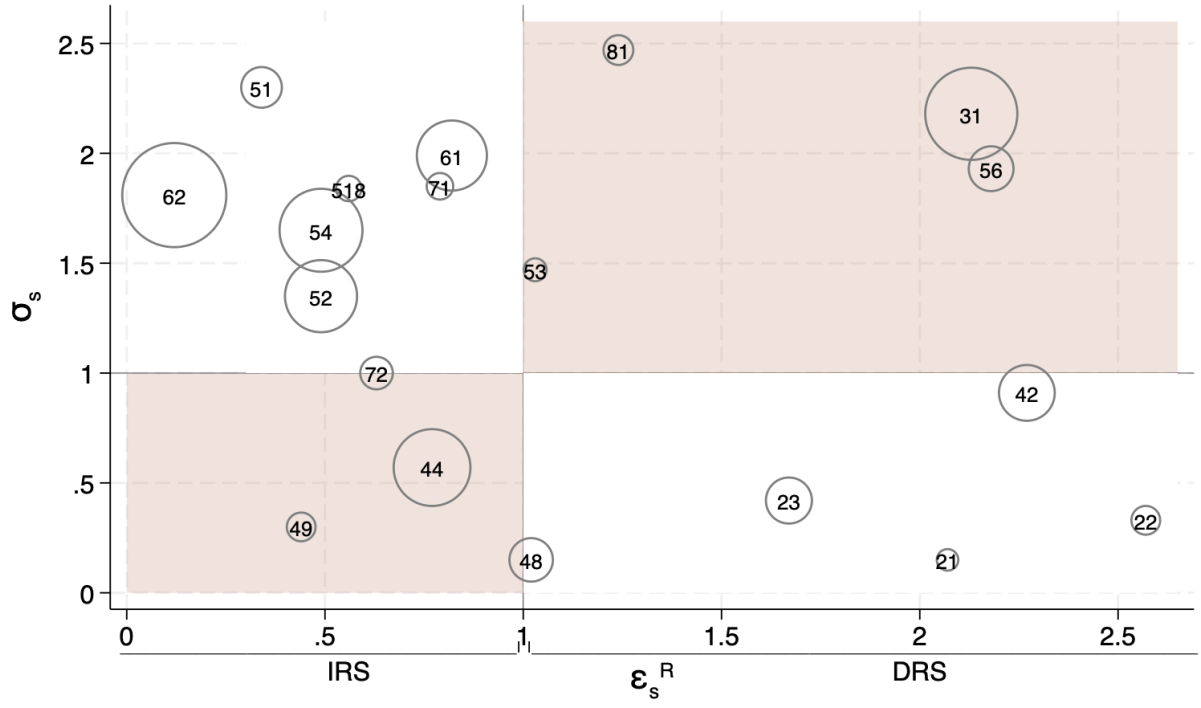
Table 2: Parameter Estimates

	Sector Title	σ		ε^R		Growth in γ_H	Skill Elasticity	Non-hom.	Skill-elastic
		Estimate	SE	Estimate	SE				
21	Mining	0.15	-	2.07	0.07	-5.90%	0.91	Y	Y
22	Utilities	0.33	0.11	2.57	0.22	40.50%	1.06	Y	Y
23	Construction	0.42	0.02	1.67	0.05	14.90%	0.38	Y	Y
31	Manufacturing	2.18	0.02	2.13	0.02	0.40%	-1.33	Y	N
42	Wholesale Trade	0.91	0.03	2.27	0.42	13.80%	0.12	Y	Y
44	Retail Trade	0.57	0.02	0.77	0.05	1.20%	-0.1	Y	N
48	Transportation	0.15	-	1.02	0.04	1.40%	0.01	N	-
49	Warehousing	0.3	0.07	0.44	0.09	6.80%	-0.4	Y	N
51	Information (exc. Tech)	2.3	0.05	0.34	0.02	27.60%	0.85	Y	Y
518	Tech (see note below)	1.84	0.12	0.56	0.1	14.20%	0.38	Y	Y
52	Finance and Insurance	1.35	0.04	0.49	0.11	16.40%	0.18	Y	Y
53	Real Estate, Rental	1.47	0.05	1.03	0.08	7.90%	-0.02	N	-
54	Prof., Sci., Tech. Serv.	1.65	0.02	0.49	0.03	11.50%	0.33	Y	Y
56	Admin., support services	1.93	0.06	2.18	0.09	1.90%	-1.1	Y	N
61	Educ. Serv.	1.99	0.07	0.82	0.03	5.10%	0.18	Y	Y
62	Health Care, Soc. Assis.	1.81	0.01	0.12	0.02	-19.10%	0.71	Y	Y
71	Arts, Ent., Rec.	1.85	0.04	0.79	0.07	-2.80%	0.17	Y	Y
72	Accomm., Food Serv.	1	-	0.63	0.03	-4.20%	-0.37	Y	N
81	Other services	2.47	0.04	1.24	0.04	3.60%	-0.35	Y	N

Note: SE denotes the standard errors of the point estimates of σ_s and ε_s^R. Growth in γ_{Hs} is the growth of the average γ_{Hfst} in the sector from 2000 to 2018. The scale elasticity is computed as $(\varepsilon_s^R - 1) \cdot (1 - \sigma_s)$ when $\sigma_s \neq 1$, and as $\varepsilon_s^R - 1$ when $\sigma_s = 1$. Non-hom. is a dummy variable that takes the value "Y" when production is non-homothetic (i.e. $\varepsilon_s^R \neq 1$), and "N" when it is homothetic. Skill-elastic takes the value of "Y' when production is skilled-elastic, "N" when it is unskilled elastic, and "-" when it is homothetic. See text for parameter configurations that determine the skill-elasticity.

1-4 starting from the upper right one and proceeding counter-clockwise. Sectors in quadrants 1 and 3 (shaded) are unskilled-elastic, while those in quadrants 2 and 4 are skill-elastic. 11 of the 17 non-homothetic sectors are skill-elastic including mining, utilities, construction, wholesale trade, information, technology, finance and insurance, professional services, education, health, and arts and entertainment. The remaining 6 non-homothetic sectors are unskilled-elastic, including manufacturing, retail trade, warehousing, administrative and support services, accommodation and food services and other services. Beyond the sign, the magnitude of the scale elasticity varies widely across sectors, with the largest (absolute) skill elasticities found in manufacturing and utilities.

Figure 2: Estimated Price and Scale Elasticities



4.5 The channel

The exogeneity of the instruments ensures that the estimated association between scale and skill reflects a causal effect. However, it does not prove that the channel through which this effect operates is linked to production non-homotheticities. A priori, the causal effect of scale on skill may operate through various channels. For example, a firm that experiences an increase in demand may face greater incentives to adopt new technologies (see Comin and Gertler, 2006). If new technologies have a particular form of skill-bias, scale would affect skill because γ_{ofst} would endogenously change as the firm adopts new technologies in response to changes in scale. Conversely, an investment bank that faces a surge in demand for wealth management services will hire more financial advisors to provide the services whose demand has increased. In this case, exogenous increases in scale lead to an increase in skill because production for financial services is non-homothetic, not because there is a change in the technologies used to provide wealth management services.

As these examples illustrate, the channels through which scale affects skill fall into two categories: Those that operate through the non-homotheticities of the production function and those

that link SBTC to the growth in scale. This observation provides an avenue to separately identify the two channels. Specifically, we can filter the potential endogeneity of γ_{ofst} to the firm's growth by permitting SBTC to differ across firms with different growth experiences as follows:

$$\ln\left(\frac{\alpha_{Hfst}}{\alpha_{Lfst}}\right) = \phi_{1s}\ln(\alpha_{Lfst}) + \phi_{2s}\ln\left(\frac{WB_{fst}}{W_{Lft}}\right) + \phi_{3s}\ln\left(\frac{W_{Hft}}{W_{Lft}}\right) + \delta_{pst} + \delta_f + \nu_{fst}. \tag{26}$$

and

$$\ln\left(\frac{\alpha_{Hfst}}{\alpha_{Lfst}}\right) = \phi_{1s}\ln\left(\frac{WB_{fst}}{\overline{W}_{fst}}\right) + \delta_{pst} + \delta_f + \nu_{fst}. \tag{27}$$

where the subindex p in δ_{pst} denotes the growth group that the firm belongs to. We measure firm growth by the actual growth in the wage bill relative to the wage of unskilled workers (i.e., WB_{fst}/W_{Lfst}). We first consider three growth groups in each sector that correspond to the bottom quartile, the middle 50% and the top quartile of the growth distribution.

In each iteration of our estimation procedure, δ_{pst} absorbs the evolution of skill in the firms in a growth quartile and reflects the sill-biased productivity of those firms. The estimated skill elasticity captures the association between exogenous variation in scale and the evolution of skill after filtering the evolution of skilled-biased productivity in firms with a similar growth experience over the period 2000-2018. Therefore, there are two differences between how δ_{pst} and non-homotheticities (ε_s^R) are identified. First, non-homotheticities only exploit exogenous variation in scale, while changes in skill captured by group-specific SBTC partly reflect the adoption choices of firms in response, among other things, to actual changes in demand. Second, δ_{pst} captures between group variation in skill, while non-homotheticities are identified out of the within group variation.

Given this identification strategy, if scale only affects skill through the differential rate of SBTC in fast vs. slow growing firms, specifications (18) and (17) will yield an estimate of $\varepsilon_s^R = 1$. Conversely, if exogenous changes in the scale of firms induce them to change their employment composition while the actual variation in demand does not induce firms to adopt technologies with a particular skill bias, our estimates will fully attribute the association between scale and skill to a skill elasticity that is different from zero, while the estimate of SBTC will not differ between fast- and slow-growing firms. Intermediate scenarios will lead to both estimates of $\varepsilon_s^R \neq 1$ and larger estimates of SBTC in the top than in the bottom growth quartiles.

Table 3 reports the estimates of the rates of SBTC in each growth group, and Table 4 com-

Table 3: Estimates of SBTC by growth quartile

			Group-Specific SBTC		
	Sector	Baseline	Bottom 25%	Middle 50%	Top 25%
21	Mining	-5.9%	-40.9%	-5.0%	27.3%
22	Utilities	40.5%	-4.3%	45.5%	78.8%
23	Construction	14.9%	11.8%	7.4%	32.5%
31	Manufacturing	0.4%	0.2%	0.3%	0.9%
42	Wholesale Trade	13.8%	-9.7%	15.5%	31.5%
44	Retail Trade	1.2%	-16.5%	3.8%	12.5%
48	Transportation	1.4%	-1.8%	-0.4%	6.5%
49	Warehousing	6.8%	15.8%	3.4%	4.4%
51	Information (except Tech)	27.6%	18.4%	21.6%	49.4%
518	Tech (see note below)	14.2%	-7.4%	26.3%	14.8%
52	Finance and Insurance	16.4%	-5.8%	20.4%	31.6%
53	Real Estate, Rental	7.9%	-6.3%	9.8%	18.1%
54	Prof, Sci, Tech. Serv.	11.5%	-8.5%	14.0%	27.6%
56	Admin, support serv.	1.9%	2.3%	1.4%	2.5%
61	Educational Services	5.1%	-4.3%	5.2%	13.4%
62	Health Care, Soc. Assis.	-19.1%	-41.3%	-15.2%	-3.8%
71	Arts, Ent., Rec.	-2.8%	-22.4%	2.6%	7.1%
72	Accomm., Food Serv.	-4.2%	-14.4%	-1.2%	7.9%
81	Other services	3.6%	0.9%	3.1%	7.7%

pares the estimates of σ_s and ε_s^R in the baseline and in specifications (18) and (17). There are some important take away from these estimates. First, there is ample variation in SBTC across growth quartiles, but, in all sectors, we find that SBTC is highest in the top growth quartile and smallest in the bottom quartile. Importantly, this pattern holds regardless of whether the sector is skill-elastic or unskilled-elastic. For example, we have found in the baseline specification that accomodation and food services is unskilled-elastic. However, the group-specific time dummies imply that the rate of SBTC in accomodation is considerably larger in firms in the top growth quartile than in firms in the bottom top quartile. This shows that specifications (26) and (27) separately identify non-homotheticities and endogenous SBTC. Second, the estimates of σ_s and ε_s^R and the skill elasticities are virtually the same as in the baseline without quartile-specific SBTC.[17] This finding provides reassurance about the exogeneity of the instruments, as the identification of the structural parameters does not reflect spurious correlation that arise from differential trends in skill for fast- vs. slow-growing firms in the sector. More importantly, the fact that the estimates of

[17]The (weighted) correlation across sectors between the skill elasticities in the baseline and when allowing for group-specific SBTC is 0.99.

Table 4: Estimates in baseline vs. group-specific SBTC

		Baseline				Group-specific SBTC			
	Sector	σ	s.d.	ϵ^R	s.d.	σ	s.d.	ϵ^R	s.d.
21	Mining	0.15	-	2.07	0.07	0.15	-	2.07	0.07
22	Utilities	0.33	0.11	2.57	0.22	0.41	0.11	2.70	0.28
23	Construction	0.42	0.02	1.67	0.05	0.42	0.02	1.71	0.05
31	Manufacturing	2.18	0.02	2.13	0.02	2.18	0.02	2.13	0.02
42	Wholesale Trade	0.91	0.03	2.27	0.42	0.90	0.02	1.74	0.26
44	Retail Trade	0.57	0.02	0.77	0.05	0.58	0.02	0.83	0.05
48	Transportation	0.15	-	1.02	0.04	0.15	-	1.04	0.04
49	Warehousing	0.30	0.07	0.44	0.09	0.29	0.07	0.49	0.09
51	Information (except Tech)	2.30	0.05	0.34	0.02	2.32	0.05	0.36	0.02
518	Tech (see note below)	1.84	0.12	0.56	0.10	1.86	0.12	0.63	0.10
52	Finance and Insurance	1.35	0.04	0.49	0.11	1.35	0.04	0.51	0.11
53	Real Estate, Rental	1.47	0.05	1.03	0.08	1.47	0.05	0.95	0.08
54	Prof., Sci., Tech. Serv.	1.65	0.02	0.49	0.03	1.66	0.02	0.52	0.03
56	Admini, support serv.	1.93	0.06	2.18	0.09	1.92	0.06	2.16	0.09
61	Educational Services	1.99	0.07	0.82	0.03	2.01	0.07	0.83	0.03
62	Health Care, Soc. Assis.	1.81	0.01	0.12	0.02	1.82	0.01	0.14	0.02
71	Arts, Ent., Rec.	1.85	0.04	0.79	0.07	1.85	0.04	0.80	0.07
72	Accomm., Food Serv.	1.00	-	0.63	0.03	1.00	-	0.51	0.05
81	Other services	2.47	0.04	1.24	0.04	2.47	0.04	1.24	0.04

the skill elasticity do not change after allowing for the possibility of growth-specific SBTC implies that the effect of exogenous changes in scale on the firm' skill that we have estimated in the baseline specifications operates through non-homotheticities in production, as opposed to reflecting endogenous skill-biased changes in technology.

Before moving on, we conduct one robustness check by providing greater flexibility to SBTC to reflect the actual growth experience of firms. We do this by using deciles, instead of quartiles, to define the growth groups p. As a result, specifications (26) and (27) now include 10 sets of time dummies, δ_{Hpst}, in each sector. Table A3 in the Appendix reports the estimates, which are completely robust to this modification. In particular, the estimates of σ_s and ε_s^R do not vary when allowing for more detailed growth groupings. This finding further strengthens our conclusion that the causal effect of scale on skill estimated in section (4.4) is due to the presence of non-homotheticities in production.

4.6 Omission of scale

The literature on the drivers of the skill premium has estimated homothetic versions of specification (20), where $\varepsilon_s^R = 1$, denying the effect of scale on skill. Next, we explore how does the omission of scale in specification (20) affects the estimates of SBTC and the elasticity of substitution between high- and low-skilled workers. To this end, we re-estimate specification (20) setting ε_s^R to 1. Table 5 reports the estimates of σ_s and the growth of γ_{Hs} between 2000 and 2018. We also report the difference between these estimates and those obtained when we do not constrain ε_s^R to be 1. This difference can be interpreted as the bias in the elasticity of substitution and in the SBTC from omitting the presence of non-homotheticities.

The omission of scale biases considerably the estimates of σ_s. In 13 out of 19 sectors, the bias in the estimate of σ_s is statistically significant. Additionally, it is often large. For example, the estimated elasticity of substitution is higher by .43 in finance and insurance and lower by 1.02 in manufacturing as a result omitting the effect of scale on skill. The bias in σ_s is positively correlated with the skill elasticity across sectors, with a weighted correlation coefficient of 0.82. That is, in sectors where scale leads to greater skill intensity, the omission of scale induces a larger increase in the estimate of σ_s. The omission of scale also results in biases in the estimates of SBTC. For example, the estimated growth in γ_H is 20% lower in utilities and 20% higher in healthcare when omitting the effect of scale on skill. However, we do not observe an association between the bias in the estimate of SBTC and the skill elasticity across sectors.

Table 5: Estimates Homothetic Model and Comparison to Baseline

		σ_s^h	SE	Growth γ_H^h	$\sigma_s^h - \sigma_s$	Growth γ_H^h / γ_H
21	Mining	1.12	0.11	5.1%	0.97**	11.0%
22	Utilities	1.62	0.06	20.7%	1.29***	-19.8%
23	Construction	0.52	0.02	15.1%	0.10***	0.2%
31	Manufacturing	1.16	0.01	7.8%	-1.02***	7.4%
42	Wholesale Trade	0.88	0.02	12.6%	-0.02	-1.2%
44	Retail Trade	0.59	0.02	0.1%	0.02	-1.2%
48	Transportation	0.06	0.04	2.0%	-0.09**	0.6%
49	Warehousing	0.07	0.08	2.6%	-0.23**	-4.2%
51	Information (exc. Tech)	1.90	0.06	12.2%	-0.4***	-15.4%
518	Tech (see note below)	2.29	0.16	5.6%	0.44**	-8.6%
52	Finance and Insurance	1.79	0.04	13.3%	0.43***	-3.1%
53	Real Estate, Rental	1.46	0.05	7.9%	-0.01	0.0%
54	Prof., Sci., Tech. Serv.	1.95	0.03	6.4%	0.31***	-5.1%
56	Admin., support serv.	2.08	0.04	5.1%	0.15**	3.3%
61	Educ. Services	2.08	0.08	2.8%	0.09	-2.3%
62	Health Care, Soc. Assis.	2.00	0.02	1.6%	0.19***	20.6%
71	Arts, Ent't, Rec.	1.93	0.04	-0.3%	0.09	2.6%
72	Accomm., Food Serv.	0.96	0.08	-9.7%	-0.04	-5.5%
81	Other services	2.27	0.03	6.2%	-0.20***	2.6%
	Correlation with skill elasticity				0.82	-0.06

4.7 Robustness and stability

We conclude our exploration of the presence of non-homotheticities in production by studying two additional issues. First, we study the role of data imputations on the estimates. Second, we study the stability of the estimates across the firm size and firm growth distributions.

Imputations and model estimates As discussed in section 3, the construction of our dataset involves imputing data on the composition of employment and wages for some establishments. To study whether the imputations affect the estimates, we re-estimate the model restricting the sample to firms where the imputations represent less than 10% of the firm's employment. Table A4 reports number of observations and estimates for this restricted sample and compares them with those in the full sample. The restricted sample is considerably smaller in all sectors. The number of firms in the restricted sample ranges from 20% (finance and insurance) to 44% (utilities) of the full sample. Despite the difference in sample size and the composition of firms in them, the estimates of σ_s and ε_s^R vary relatively little between the two samples. In all instances, we find that the pair $\{\varepsilon_s^R, \sigma^s\}$ fall in the same quadrant as in the full-sample estimation. As a result, the skill elasticities

are very similar in both samples. Specifically, the weighted correlation between the skill elasticities in both samples across sectors is 0.99. Therefore, we conclude that the data imputations made to assemble the dataset do not drive the estimates of the production function and in particular the finding of systematic non-homothetities in production.

Stability of estimates Before studying the implications of the non-homothetic production functions, we explore the stability of the estimates across the distributions of firms. In particular, we consider two relevant dimensions to rank firms. First, their initial size measured by the firm's wage bill to unskilled wage ratio in year 2000. Second, the firm's growth from 2000 to 2018, measure by the growth rate of the wage bill relative to the wage of unskilled workers. For both metrics, we classify firms in three groups that correspond to the bottom quartile, the middle 50% and the top quartile of the distributions of firm-size and firm growth, respectively. We then re-estimate the model allowing the scale elasticity parameter, ε_s^R to vary across the three size groups. The elasticity of substitution, σ_s and the time dummies are the same across the three groups.[18] Table 6 reports the estimates when ε_s^R varies across initial size quartiles, and Table 7 when it varies across growth quartiles.In all sectors, we find that the estimates of ε_s^R virtually do not vary across the three size groups, and their differences are not statistically significant at conventional levels. Given the stability of non-homotheticities across firm sizes within sectors, in the rest of the paper we use the same value of ε_s^R for all firms in a given sector.

5 Drivers of the number of efficiency units per worker

Proposition 1 presents the model predictions for the drivers of the gap between the number of efficiency units and the number of employees in a firm. In this subsection, we study whether these predictions are consistent with the data. To this end, we use data on the number of workers in high- and low-skilled occupations, the estimates of the production function, and equation (2) to compute the number of efficiency units in the firm. Using as a dependent variable the ratio of the number of efficiency units to the number of workers in the firm, we construct the data counterparts of the comparative statics in Proposition 1 by estimating the slope coefficients in the following regressions:

[18] The estimates are robust to allowing time dummies to be group-specific.

Table 6: Stability of non-homotheticities across the firm-size distribution

				Bottom Q.		Mid. 50%		Top Q.	
		σ	s.e.	ε_s^R	s.e.	ε_s^R	s.e.	ε_s^R	s.e.
21	Mining	0.15	-	2.10	0.08	2.07	0.07	2.08	0.07
22	Utilities	0.32	0.11	2.48	0.21	2.48	0.21	2.53	0.21
23	Construction	0.42	0.02	1.81	0.06	1.79	0.05	1.74	0.05
31	Manufacturing	2.18	0.02	2.11	0.02	2.12	0.02	2.12	0.02
42	Wholesale Trade	0.88	0.02	1.60	0.22	1.66	0.22	1.70	0.22
44	Retail Trade	0.57	0.02	0.97	0.05	0.95	0.05	0.84	0.05
48	Transportation	0.15	-	1.07	0.04	1.05	0.04	1.04	0.04
49	Warehousing	0.28	0.08	0.59	0.08	0.56	0.08	0.49	0.09
51	Information (exc. Tech)	2.31	0.05	0.36	0.02	0.34	0.02	0.35	0.02
518	Tech (see note below)	1.88	0.13	0.70	0.10	0.66	0.10	0.63	0.10
52	Finance and Insurance	1.37	0.05	0.60	0.10	0.57	0.10	0.55	0.11
53	Real Estate, Rental	1.43	0.05	0.75	0.09	0.82	0.08	0.89	0.08
54	Prof., Sci., Tech. Serv.	1.64	0.02	0.55	0.03	0.52	0.04	0.51	0.04
56	Admin., support serv.	1.90	0.06	2.05	0.09	2.09	0.09	2.13	0.09
61	Educ. Services	1.98	0.07	0.81	0.04	0.81	0.04	0.82	0.04
62	Health Care, Soc. Assis.	1.81	0.01	0.14	0.02	0.13	0.02	0.12	0.02
71	Arts, Ent't, Rec.	1.84	0.04	0.78	0.07	0.79	0.07	0.78	0.07
72	Accomm., Food Serv.	1.00	-	0.79	0.03	0.76	0.03	0.67	0.03
81	Other services	2.47	0.04	1.22	0.04	1.21	0.04	1.23	0.04

Notes: Bottom Q. and Top Q. stand for Bottom Quartile and Top Quartile of the sectoral firm size distribution.

$$\ln\left(\frac{L_{fst}}{\overline{L}_{fst}}\right) = \kappa_0 + \kappa_1 \cdot \ln\left(\frac{W_{Lft}}{W_{Hft}}\right) + u_{ft} \tag{28}$$

$$\ln\left(\frac{L_{fst}}{\overline{L}_{fst}}\right) = \rho_{0,s} + \rho_{1,s} \cdot \ln\left(L_{fst}\right) + u_{ft} \tag{29}$$

In the model, the sign of the coefficient κ_1 is equal to the sign of the derivative $\frac{\partial L_{fst}/\overline{L}_{fst}}{\partial W_{Lt}/W_{Ht}}$. Part A of proposition 1 states that this should be positive. Similarly, the model also implies that the sign of the coefficient ρ_1 is equal to the sign of $(1-\varepsilon^R)$. Because ε^R is sector-specific, we index $\rho_{1,s}$ by s.

The point estimate for κ_1 is 0.2 (s.e. 0.005). The estimates for $\rho_{1,s}$ are reported in Table A2 in the Appendix. In each case other than for transportation, the sign of the point estimate is the same as the sign of $(1-\varepsilon^R)$. Furthermore, the (weighted) correlation between the point estimates of $\rho_{1,s}$ and $(1-\varepsilon^R)$ is 0.9. These estimates confirm the model predictions about the gap between the number of efficiency units and the number of employees. The conformity between the model prediction and the actual estimates of κ_1 and $\rho_{1,s}$ supports the structural interpretation of the

Table 7: Stability of non-homotheticities across the firm-growth distribution

				Bottom Q.		Mid. 50%		Top Q.	
	Sector	σ	s.e.	ε_s^R	s.e.	ε_s^R	s.e.	ε_s^R	s.e.
21	Mining	0.15	-	1.09	0.03	1.21	0.03	1.18	0.03
22	Utilities	1.48	0.06	0.84	0.03	0.73	0.05	0.81	0.04
23	Construction	0.53	0.02	0.99	0.02	1.06	0.02	0.96	0.02
31	Manufacturing	1.12	0.01	0.88	0.03	0.45	0.07	0.84	0.04
42	Wholesale Trade	0.89	0.04	3.16	0.73	4.95	1.35	2.04	0.36
44	Retail Trade	0.58	0.02	0.82	0.02	1.10	0.02	0.70	0.02
48	Transportation	0.15	-	0.95	0.01	0.97	0.01	1.00	0.01
49	Warehousing	0.11	0.08	0.88	0.03	0.93	0.03	0.95	0.02
51	Information (exc.Tech)	2.01	0.06	0.96	0.01	0.73	0.02	0.95	0.01
52	Tech (see note below)	1.58	0.04	0.81	0.02	0.70	0.03	0.87	0.01
53	Finance and Insurance	1.48	0.05	1.00	0.03	1.03	0.04	1.16	0.04
54	Real Estate, Rental	1.81	0.02	0.82	0.01	0.63	0.01	0.86	0.01
56	Prof., Sci., Tech. Serv.	2.04	0.04	1.09	0.01	1.11	0.01	1.11	0.01
61	Admin., support serv.	2.08	0.07	0.96	0.01	0.92	0.01	1.00	0.01
62	Educ. Services	1.99	0.01	0.90	0.00	0.82	0.01	0.89	0.00
71	Health C., Soc. Assist	1.92	0.04	0.93	0.02	0.96	0.02	0.98	0.02
72	Arts, Ent., Rec.	1.00	-	0.76	0.02	0.65	0.02	0.71	0.02
81	Accom., Food Serv.	2.26	0.03	1.00	0.01	0.94	0.01	1.04	0.01
518	Other services	2.06	0.13	0.90	0.03	0.74	0.05	0.88	0.03

Notes: Bottom Q. and Top Q. stand for Bottom Quartile and Top Quartile of the sectoral firm growth distribution.

estimates of the production function parameters in section 4.4.

6 Drivers of skill

Next, we study the role of non-homotheticities for the skill intensity of firms. We first focus on the level of skill and conduct a variance decomposition to study the contribution of scale, relative wages and SBTC to the variance of $log(\alpha_{Hfst})$. Then, we shift our focus to the change in skill. For each firm, we use the estimated model to compute the predicted change in skill between 2000 and 2018 and compare the model predictions to the data. Finally, we study the role of each channel for the evolution of skill by using model counterfactuals.

Table 8: Decomposition of the Variance of $\ln(\alpha_{f,s,t})$

NAICS	Sector	Var $\left[\ln(\alpha_{H,f,s,t})\right]$	R^2	Contribution to Var $\left[\ln(\alpha_{H,f,s,t})\right]$		
				SBTC	Rel. wages	Scale
21	Mining	0.71	52%	82.9%	-19.9%	-11.1%
22	Utilities	0.31	37%	50.5%	-17.7%	4.4%
23	Construction	0.49	31%	34.5%	1.6%	-4.8%
31	Manufacturing	0.43	33%	1.4%	13.7%	17.6%
42	Wholesale Trade	0.45	32%	36.7%	-2.8%	-2.1%
44	Retail Trade	0.57	43%	34.1%	5.0%	3.6%
48	Transportation	0.68	57%	40.8%	16.2%	0.0%
49	Warehousing	0.52	48%	14.8%	38.5%	-5.1%
51	Information (except Tech)	0.35	43%	43.1%	-34.6%	34.7%
518	Tech (see note below)	0.18	38%	40.9%	-62.4%	59.7%
52	Finance and Insurance	0.27	42%	47.4%	-26.7%	21.0%
53	Real Estate, Rental & Leasing	0.71	35%	31.6%	3.2%	0.3%
54	Prof., Sci., Tech. Serv. & Mg't	0.21	37%	43.2%	-51.5%	45.5%
56	Admin & Support serv.	0.97	58%	2.6%	17.2%	38.3%
61	Educational Services	0.18	19%	28.3%	-31.9%	22.6%
62	Health Care & Social Assist.	0.51	61%	51.2%	-80.5%	90.1%
71	Arts, Ent'ment & Recreation	0.72	55%	53.3%	-0.6%	2.2%
72	Accommodation & Food Serv.	0.85	-33%	16.5%	0.0%	-49.6%
81	Other services	0.79	64%	13.9%	41.0%	8.9%
	Weighted Average		41.8%	34.5%	-24.3%	30.5%

Note: Tech sector includes Software publishers, Data processing, Hosting, and Related Services, and Internet Publishing and Broadcasting and Web Search

6.1 Cross-sectional variance decomposition

We conduct a variance decomposition of the (log) wage bill share of high-skilled workers ($\ln(\alpha_{Hfst})$). We first compute the three terms in expression (9) for each firm.[19] The contribution of a term in (9) to the variance of (log) skill is given by the covariance between the log of the term and $\ln(\alpha_{Hfst})$ divided by the variance of $\ln(\alpha_{Hfst})$.

The first column of Table 8 reports the variance of the (log) of skill across the firms over the seven observations that make up our sample for each sector. Columns 3 through 5 report the contribution of each term, for each sector. Column 2 reports the sum of columns 3 through 5 which is the fraction of the variance of (log) skill explained by the model.

On average, the model accounts for 42% of the variance of $\ln(\alpha_{Hfst})$. Both SBTC and Scale contribute significantly to this explanatory power. SBTC accounts for 34.5% of the variance in

[19] We first construct γ_{Hfst} and γ_{Lfst} using the firm random effects and the sector-specific time-dummies. Together with the estimates of σ_s and ε_s^R and the data on the number of high- and low-skilled employees in the firm, we use expression (2) to compute L_{fst}. The scale term is $(\varepsilon_s^R - 1)(1 - \sigma_s) \cdot \ln(L_{fst})$. Given the wage bill in the firm, we use expression (5) to construct $\widehat{W}_{fst}\bar{\varepsilon}_{fst}$. The relative wage term is $(\sigma_s - 1) \cdot \ln\left(\frac{\widehat{W}_{fst}/\bar{\varepsilon}^L_{fst}}{W_{Hft}}\right)$. The SBTC is $\ln(\gamma_{Hfst})$.

skill, while scale accounts for 31.4%. Relative wages, have on average a negative contribution (-24.3%). The relevance of the three channels varies across sectors. Scale is the main factor in manufacturing, technology (NAICS 518), professional services, administrative and support services and healthcare. Relative wages are the main factor in other services. SBTC is the main driver in the rest of the sectors. It's worth noting that much of the variation in the SBTC reflects the firm random effects. On average, the time dummies account for only 0.5% of the variance in skill.[20]

6.2 The evolution of skill

Next, we explore the model's predictions for the evolution of skill between 2000 and 2018. We use data on firm's employment and wages by occupation together with the estimated model parameters to calculate the predicted change in skill for each firm. With these firm-level predictions we conduct two different exercises. First, we compute the average predicted change in skill across all the firms in each sector and compare them to the data. Second, we investigate whether the model captures the different changes in skill observed between fast- and slow-growing firms within each sector. After analyzing how the model fits the data, we study the robustness of the findings to allowing SBTC to differ across firms in different quartiles of the growth distribution. We conclude this section by examining the role played by the different channels in the predicted change in skill through a counterfactual analysis.

Baseline estimated model

For each of the 19 2-digit sectors, we study the change in skill in the data and the model. Our first step consists in constructing the model predicted level of skill for each firm and year as follows. First, we use the estimated values of γ_{Hfst} and γ_{Lfst}, σ_s and ε_s^R together with the number of employees in high- and low-skilled occupations and expression (2) to compute the firm' scale measured by the number of efficiency units (L_{fst}). Then, we use information on firm wages, the wage bill and equation (6) to compute the wage bill share of high skill workers. Finally, we construct the average change in skill across the relevant sample of firms as a simple average of the firm-level changes. We consider various subsamples of the firm growth distribution to study the relationship between skill change and firm growth, in each sector.

[20] Table A3 in the Appendix reports the variance decomposition excluding the firm random effect (ρ_f). This directly affects the variation in the SBTC term, which now is only driven by time dummies, and also the other two terms, as L_{fst} and $\widehat{W}_{fst}$ depend on γ_{Hfst} and γ_{Lfst}. The average R^2 is still 19%, Scale accounts for 24.2% of the variance in skill, and SBTC for 0.5%.

Columns 1 and 2 of Table 9 report the (average) increase in skill across all the firms in each sector. At the bottom of the table, we report the weighted average of the changes in skill across all sectors, and the cross-sector weighted correlation between the actual and the predicted average changes in skill. The sectoral weights are given by the sector's share in the economy-wide wage bill. The model predictions for the average change in skill from 2000 to 2018 are very close to the changes observed in the data. For example, in professional services, the model predicts an average increase by 7.45 percentage points while in the data we observe an average increase by 7.27 percentage points. The (weighted) average across sectors of the average predicted increase in skill is 3.14 percentage points while in the data it is 3.31 percentage points. The model does a good job in explaining the variation across sectors in the average change in skill. The (weighted) correlation between the average increases in skill in the model and the data across sectors is 0.98.

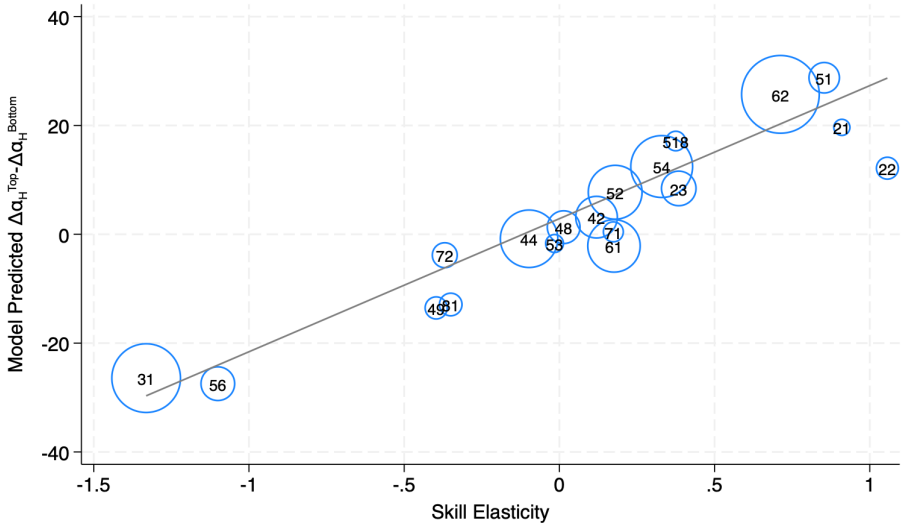
Table 9: Increase (Percentage points) in skill-intensity from 2000-2018

		All Firms		Middle 50%		Bottom Quartile		Top Quartile		Top-Bottom	
		Data	Model	Data	Model	Data	Model	Data	Model	Data	Model
21	Mining	0.97	1.61	1.07	0.60	-3.55	-7.16	5.32	12.49	8.87	19.64
22	Utilities	6.07	5.43	6.49	4.85	0.74	-0.07	10.54	12.08	9.80	12.16
23	Construction	2.86	2.25	1.98	2.10	2.24	-1.81	5.27	6.62	3.03	8.43
31	Manufacturing	2.72	2.52	1.84	3.22	1.31	15.04	5.87	-11.39	4.56	-26.42
42	Wholesale Trade	3.60	3.79	4.02	3.81	-1.89	2.20	8.23	5.35	10.12	3.15
44	Retail Trade	0.42	0.21	1.06	0.32	-3.23	0.50	2.78	-0.32	6.01	-0.82
48	Transportation	0.34	-0.16	0.61	0.15	-0.52	-1.13	0.66	0.18	1.18	1.31
49	Warehousing	-0.46	0.42	-0.88	-0.42	1.34	8.00	-1.42	-5.53	-2.76	-13.54
51	Information (ex. Tech)	7.18	5.04	5.16	5.53	5.29	-9.83	13.11	18.94	7.82	28.77
518	Tech (see note below)	8.16	7.56	11.01	8.70	1.63	-2.15	8.99	14.99	7.36	17.14
52	Finance and Insurance	8.70	9.19	9.82	9.04	0.59	5.48	14.55	13.21	13.95	7.73
53	Real Estate, Rental	3.21	2.60	3.98	2.82	-1.12	3.23	5.99	1.54	7.11	-1.69
54	Prof., Sci., Tech. Serv.	6.83	7.26	7.27	7.45	0.14	0.85	12.63	13.30	12.49	12.45
56	Admin., support serv.	3.04	3.01	1.60	2.32	4.73	17.43	4.21	-10.02	-0.52	-27.45
61	Educational Services	3.70	4.24	3.94	4.67	-0.49	4.88	7.42	2.75	7.91	-2.13
62	Health Care, Soc. Assis.	0.65	-0.10	1.46	0.26	-5.88	-13.32	5.55	12.40	11.43	25.72
71	Arts, Ente., Rec.	0.65	0.66	1.91	0.29	-3.73	0.80	2.51	1.25	6.24	0.45
72	Accomm., Food Serv.	-0.05	-0.84	0.22	-0.84	-0.33	1.08	-0.31	-2.76	0.02	-3.83
81	Other services	4.33	4.65	3.57	4.11	1.10	11.63	9.09	-1.27	7.99	-12.89
	Weighted Average	3.31	3.16	3.54	3.35	-1.07	0.99	7.21	4.91	8.28	3.93
	Weighted Correlation	0.98		0.95		0.7		0.57		0.65	

Note: Weighted correlation reports correlation between data and model across sectors using the sector's wage bill share in the economy as weights. Note: Mining is short for "Mining, Quarrying, and Oil and Gas Extraction "; Admin. & Support, Waste Mg't is short for "Administrative and Support and Waste Mananagment and Remediation Services."

Next, we investigate the evolution of skill within sectors. We sort firms into three groups based on the growth of their wage bill relative to the wage of unskilled workers (WB_{fst}/W_{Lfst}) between 2000 and 2018. For each sector, we compute the (simple) average of the change in skill for the firms in the middle 50% of the distribution (columns 3 and 4), in the bottom quartile (columns 5 and 6), and in the top quartile (columns 7 and 8) in the data and in the model.

Figure 3: Change in skill intensity top vs bottom



We start by analyzing the average evolution of skill for the firms in the middle 50% of the growth distribution which are very similar to the average evolution for the entire sample both in the data and in the model.[21] In most sectors, the average change in skill predicted by the model is very close to the data. For example, in professional services the model predicts an increase by 7.45 percentage points and we observe an increase by 7.27 percentage points. On average, skill increases by 3.35 percentage points (vs. 3.54 in the data). The (weighted) correlation between the actual and predicted average change in skill across sectors for firms in the middle 50% of the growth distribution is 0.95.

We now examine the evolution of skill in firms at the top and bottom quartiles of the growth distribution. A key finding is that skill changes differ significantly between fast- and slow-growing firms. For example, in professional services, firms in the bottom quartile saw an average skill increase of just 0.14 percentage points, whereas firms in the top quartile experienced an increase of 12.63 percentage points. This pattern generally holds across sectors, with the exceptions of warehousing, administrative and support services, and accommodation and food services.On average, firms in the bottom quartile of the growth distribution have experienced a decline in skill by 1.07 percentage points, whereas firms in the top quartile have seen an increase of 7.21 percentage points.

[21]The (weighted) average across sectors of the change in skill has been 3.54 percentage points for the middle 50% and 3.31 for the entire sample, and the weighted correlation of the change in skill for the middle 50% and the entire sample across sectors is 0.96.

The model captures the different evolution of skill for firms in the top and bottom quartiles of growth, though not as closely as for the entire sample or across firms in the middle 50% of the growth distribution. For example, in professional services, the model predicts an increase in skill for the firms at the bottom growth quartile by 0.85 percentage points (vs. 0.14 in the data) while for the firms at the top growth quartile it predicts an increase by 13.3 percentage points (vs. 12.63 in the data). Across sectors, the weighted average of the change in skill in the model at the bottom growth quartile is 0.99 percentage points (vs. -1.07 in the data), and at the top growth quartile it is 4.91 percentage points (vs. 7.21 in the data). The (weighted) correlation between the actual and predicted average change in skill across sectors is 0.7 for the bottom quartile and 0.57 for the top growth quartile. These correlations suggest the model captures well the cross-sector variation in skill changes for fast and slow-growing firms relatively. However, in some sectors there are significant deviations from the data. For example, in all six unskilled elastic sectors, the model predicts larger increases in skill for slow-growing firms than for fast-growing ones. However, we observe this pattern in only two sectors (warehousing and administrative and support services).

To better assess the model's ability to account for the differential change in skill in fast-growing and slow-growing firms in each sector, we subtract the average increase in skill for firms in the bottom growth quartile (columns 5 and 6) from the average increase for firms in the top growth quartile (columns 7 and 8), and report the results in columns 9 and 10 of Table 9. The model accounts for approximately half of the differential increase in skill between fast-growing and slow-growing firms. Specifically, the model explains a 3.93 percentage point increase, compared to the 8.28 percentage points observed in the data. The (weighted) correlation between the differential increase in skill (top minus bottom quartiles) in the model and the data across sectors is 0.65. Despite this relatively high correlation, in four sectors—manufacturing, retail trade, real estate and other services—the model predicts a larger increase in skill for slow-than fast-growing firms, while the data shows the opposite.

Below, we use counterfactuals to study the channels that drive the differential change in skill in fast- vs. slow growing firms in the model. However, we can start to explore the role of scale by studying the association across sectors between the skill elasticity and the differential change in skill in fast vs. slow-growing firms. Figure 3 presents a scatter plot of these two variables. There is a very strong association between them, with a weighted correlation of 0.96. This finding suggests that non-homotheticities drive the model's within-sector variation in the evolution of skill between fast and slow-growing firms. Additionally, the weighted correlation between the

skill elasticity and the actual differential change in skill in the data in fast vs. slow growing firms across sectors is 0.6. Therefore, non-homotheticities are important to explain the observed within sector variation in the evolution of skill both in our model and in the data.

Group-specific SBTC

To study the discrepancies between model and data in the differential growth of skill between fast- and slow-growing firms, we redo the analysis using the version of the estimated model with quartile-specific SBTC. Table 10 reports the average increase in skill between 20000 and 2018 in each subsample.

The main findings are as follows. The model performance for the entire and the middle 50% samples is robust to introducing quartile-specific SBTC. However, the introduction of quartile-specific SBTC enhances the model capacity to produce a differential increase in skill in fast-vs. slow-growing firms. The (weighted) average differential change in skill is 12.3 percentage points whereas in the data it is 8.28 percentage points. The (weighted) correlation between the differential change in skill in model and data across sectors increases to 0.76. In particular, the model accounts better for the differential increase in skill in fast- vs. slow-growing firms in retail trade, real estate and other services. This is the case because in these sectors, we have estimated a larger SBTC for fast- than for slow-growing firms. However, in manufacturing and other services, SBTC is similar for fast and slow growing firms. As a result, the model still predicts larger increases in skill for slow-growing firms, whereas the data shows the opposite.

3-digit disaggregation

Next, we explore whether heterogeneity in production functions across 3-digit sectors may help explain the differential change in skill in fast- vs. slow-growing firms in manufacturing and other services. For each 3-digit sector in manufacturing and other services,[22] we construct three growth groups following the same procedure as above, and estimate the version of the model with quartile-specific SBTC (i.e. specifications 26 and 27). Note that, this exercise extends the previous estimation by allowing the parameters σ_s and ε_s^R to vary across the 3-digit sectors of manufacturing and other services.

Table A6 contains the model estimates and the average across 3-digits for the skill-elasticity and quartile-specific SBTC, where 3-digit sectors are weighted by the share of observations in

[22] In manufacturing there are 18 3-digit sectors, while in other services there are 8.

the 2-digit sector. We find substantial heterogeneity within manufacturing and other services. 16 out of 18 3-digit manufacturing sectors are unskill-elastic while petroleum and coal products and computer and electronic products are skill-elastic. However, the average skill-elasticity is higher than in the baseline (-1 vs. -1.33). In other services, we find that half of the 3-digit sectors are skilled-elastic while the other half are unskilled elastic. In this case, the average skill-elasticity is slightly smaller than in the baseline (-.4 vs. -.35). We also find significant heterogeneity in the rates of SBTC across growth quartiles and 3-digit sectors. However, in both manufacturing and other services, on average, there is a larger difference in the rate of SBTC between fast-and slow-growing firms than in estimation at the 2-digit level.

Table A7 reports the model prediction for the average across the three digit sectors of the differential change in skill in fast-vs. slow-growing firms for each of the three versions of the model we have estimated. Recognizing the heterogeneity of the production functions and SBTC across 3-digit sectors improves the model account of the differential change in skill in fast- vs. slow-growing firms in manufacturing and other services. In particular, the gap (between model and data) decreases from 27 to 18 percentage points in manufacturing and from 13 to 5 percentage points in other services. The improvement of the fit results from both the higher skill-elasticity (in manufacturing) and the steeper pattern of SBTC across the growth quartiles (in both sectors).

Counterfactuals

We conclude this section by conducting counterfactual analyses to explore the role played by the three mechanisms – SBTC, relative wages and scale – in the evolution of skill predicted by the model. The starting point are the values of $\gamma_{H,f,st}$, $\gamma_{L,f,st}$, $\widehat{W}_{fst}$, W_{Hfst}, W_{Lfst}, and L_{fst} computed for each firm in 2000 and 2018 in section 6.2. Given those, we compute the counterfactual levels of skill in 2018 if only one of the three forces changed to the 2018 level keeping the other two at their level in 2000.[23]

Table 11 reports the average increase in skill across all firms in each sector for each of the three counterfactuals. There are two main findings. Scale and SBTC contribute by a similar amount to the (weighted) average predicted increase in skill between 2000 and 2018 by the model. In both cases, the increase is around 1.5 percentage points while the overall increase in the model is 3.14 percentage points. Changes in the average skill level induced by SBTC are more correlated across

[23]We compute the predicted level of skill first by computing the predicted levels of employment in each occupation using equation (6), and then computing skill as the wage bill share of high skill workers combining the predicted levels of employment and the relevant wages.

sectors with the overall changes predicted by the model than changes induced by scale (0.86 vs. 0.28). These findings suggest that cross-sectoral differences in SBTC, and not in the average change in scale, are the main driver of the variation across sectors in the change in skill predicted by the model. However, both scale and SBTC are equally important to understand the aggregate change in skill predicted by the model.

Next, we study the drivers of the model predictions for the evolution of skill within sectors by calculating the average change in skill for the firms in the bottom and top quartiles of the growth distribution under each of the three counterfactuals. (See Table 12.)[24] The main finding is that scale is the only force driving the differential change in skill produced by the model between fast- and slow-growing firms. Let's use professional services as an example that reflects the broader patterns. Scale reduces skill among the firms in the bottom quartile of the growth distribution by 4.75 percentage points and increases it among the firms in the top quartile by 11.28 percentage points. Therefore, it contributes to increasing the gap in skill between fast and slow growing firms by 16.03 percentage points. This differential increase reflects the entire differential increase predicted by the model which is 12.45 percentage points. In contrast, SBTC produces no differential increase in skill for fast vs. slow growing firms (-0.13), and relative wages produce a larger increase for slow than for fast growing firms (-2.48). Across sectors, scale produces an average decrease in skill by 1.61 percentage points for the firms at the bottom growth quartile and an average increase by 4.1 for the firms at the top growth quartile. As a result, scale produces an average differential increase in skill between fast and slow growing firms of 5.71 percentage points, larger than the average differential produced by the model of 3.93 percentage points. SBTC produces an average differential increase for fast vs. slow growth firms of 0.12 percentage points and relative wages a differential reduction by 1.71 percentage points.

In addition to driving the average differential change in skill in fast- vs. slow-growing firms predicted by the model, scale is responsible for its variation across sector. The correlation between the differential change in skill predicted by the model and under the scale counterfactual across sectors is 0.99. In contrast, the average changes in skill under the SBTC and the relative wages counterfactual are essentially uncorrelated across sectors with the average change in skill predicted by the model.

Furthermore, scale is not only key to account for the variation in differential change in skill

[24] **??** in the Appendix contains the analysis for the firms in the middle 50% of the growth distribution. The findings are similar to those for the aggregate of all firms.

Table 10: Increase (Percentage points) in skill-intensity from 2000-2018, Group-specific SBTC

		All Firms		Middle 50%		Bottom Quartile		Top Quartile		Top-Bottom	
		Data	Model	Data	Model	Data	Model	Data	Model	Data	Model
21	Mining	0.97	1.68	1.07	0.68	-3.55	-10.82	5.32	16.31	8.87	27.13
22	Utilities	6.07	5.60	6.49	5.42	0.74	-4.73	10.54	16.28	9.80	21.01
23	Construction	2.86	2.34	1.98	1.15	2.24	-2.42	5.27	9.50	3.03	11.92
31	Manufacturing	2.72	2.35	1.84	2.16	1.31	13.70	5.87	-8.61	4.56	-22.31
42	Wholesale Trade	3.60	3.66	4.02	4.04	-1.89	-3.57	8.23	10.15	10.12	13.72
44	Retail Trade	0.42	0.21	1.06	0.87	-3.23	-2.79	2.78	1.91	6.01	4.70
48	Transportation	0.34	-0.15	0.61	-0.04	-0.52	-1.67	0.66	1.13	1.18	2.80
49	Warehousing	-0.46	0.04	-0.88	-1.63	1.34	8.41	-1.42	-5.03	-2.76	-13.44
51	Information (ex. Tech)	7.18	5.11	5.16	3.99	5.29	-11.86	13.11	24.30	7.82	36.16
518	Tech (see note below)	8.16	8.00	11.01	12.99	1.63	-6.97	8.99	13.00	7.36	19.97
52	Finance and Insurance	8.70	9.26	9.82	10.41	0.59	-3.12	14.55	19.32	13.95	22.43
53	Real Estate, Rental	3.21	2.51	3.98	3.19	-1.12	-1.06	5.99	4.73	7.11	5.79
54	Prof., Sci., Tech. Serv.	6.83	7.00	7.27	8.25	0.14	-6.16	12.63	17.65	12.49	23.81
56	Admin., support serv.	3.04	2.71	1.60	0.53	4.73	18.71	4.21	-8.92	-0.52	-27.62
61	Educational Services	3.70	4.14	3.94	4.75	-0.49	0.11	7.42	6.96	7.91	6.85
62	Health Care, Soc. Assis.	0.65	-0.15	1.46	1.36	-5.88	-19.81	5.55	16.47	11.43	36.28
71	Arts, Ent., Rec.	0.65	0.68	1.91	1.44	-3.73	-4.05	2.51	3.89	6.24	7.94
72	Accomm., Food Serv.	-0.05	-0.82	0.22	-0.44	-0.33	-0.64	-0.31	-1.79	0.02	-1.15
81	Other services	4.33	4.60	3.57	3.41	1.10	8.24	9.09	3.32	7.99	-4.91
	Weighted Average	3.31	3.07	3.54	3.65	-1.07	-3.50	7.21	8.53	8.28	12.03
	Weighted Correlation	0.99		0.99		0.72		0.63		0.76	

Note: Weighted correlation reports correlation between data and model across sectors using the sector's wage bill share in the economy as weights.

predicted by the model across sectors, it is also critical to account for the variation in the differential change in skill observed in the data. The (weighted) correlation across sectors between the differential change in skill between fast- and slow-growth firms under the scale counterfactual and in the data is 0.68.

In the Appendix, we have explored the robustness of these findings to adopting the version of the model with quantile-specific SBTC. (See Table A8 and Table A9)[25] The main conclusions are robust to using this specification. In particular, scale continues to account for about half of the average increase in skill across firms. It now accounts for an average differential increase in skill in fast- vs. slow-growing firms of 8.3 percentage points (vs. 5.7 in the baseline). The main difference with the baseline is that SBTC accounts for an average differential increase in skill in fast- vs. slow-growing firms of 7.25 (vs. 0 in the baseline). However, the differential increases induced by scale are more strongly correlated with the increases produced by the model than those induced by SBTC (0.97 vs. .52).

From the counterfactual analysis, we conclude that scale is at least as important as SBTC to

[25] In manufacturing and other services, we conduct the counterfactual exercises at the 3-digit level and then aggregate them using the share in the total number of firms in each sector.

Table 11: Counterfactual increases in skill, All firms Baseline

			Channels		
Sector		Model	SBTC	Wages	Scale
21	Mining	1.61	-0.83	-0.16	2.67
22	Utilities	5.43	6.54	-0.37	-0.61
23	Construction	2.25	2.26	-0.46	0.48
31	Manufacturing	2.52	1.56	0.82	0.15
42	Wholesale Trade	3.79	3.65	-0.02	0.15
44	Retail Trade	0.21	0.23	0.21	-0.22
48	Transportation	-0.16	0.19	-0.38	0.03
49	Warehousing	0.42	2.36	-0.71	-1.08
51	Information (ex. Tech)	5.04	4.83	0.79	-0.53
518	Tech (see note below)	7.56	3.78	0.04	4.27
52	Finance and Insurance	9.19	5.86	0.21	3.20
53	Real Estate, Rental	2.60	2.15	0.46	-0.04
54	Prof., Sci., Tech. Serv.	7.26	3.55	0.50	3.51
56	Admin., support serv.	3.01	3.22	0.35	-0.60
61	Educational Services	4.24	2.60	0.85	0.92
62	Health Care, Soc. Assis.	-0.10	-3.32	-0.39	3.60
71	Arts, Ent., Rec.	0.66	-0.64	0.61	0.71
72	Accomm., Food Serv.	-0.84	-0.59	0.00	-0.27
81	Other services	4.65	4.22	0.91	-0.55
	Weighted Average	3.16	1.49	0.22	1.51
	Weighted Correlation		0.86	0.44	0.28

Table 12: Counterfactual changes in skill by growth quartile, baseline

		Bottom 25%				Top 25%				Top−Bottom			
		Model	SBTC	Wages	Scale	Model	SBTC	Wages	Scale	Model	SBTC	Wages	Scale
21	Mining	-7.16	-0.98	-0.63	-5.65	12.49	-0.93	0.20	13.27	19.64	0.05	0.84	18.92
22	Utilities	-0.07	7.44	-1.20	-5.44	12.08	7.83	-0.10	3.83	12.16	0.39	1.10	9.27
23	Construction	-1.81	2.55	-0.89	-3.17	6.62	2.30	-0.20	4.23	8.43	-0.26	0.70	7.40
31	Manufacturing	15.04	1.06	1.47	11.81	-11.39	1.51	-0.34	-12.08	-26.42	0.45	-1.81	-23.89
42	Wholesale Trade	2.20	3.75	-0.19	-1.26	5.35	3.69	0.10	1.45	3.15	-0.05	0.29	2.70
44	Retail Trade	0.50	0.22	-0.47	0.76	-0.32	0.21	0.65	-1.15	-0.82	-0.01	1.12	-1.92
48	Transportation	-1.13	0.19	-1.20	-0.12	0.18	0.19	-0.18	0.17	1.31	-0.01	1.02	0.29
49	Warehousing	8.00	1.76	-0.67	6.45	-5.53	1.91	0.15	-6.89	-13.54	0.15	0.83	-13.33
51	Information (ex. Tech)	-9.83	4.21	2.48	-16.34	18.94	4.86	-1.76	15.98	28.77	0.66	-4.25	32.31
518	Tech (see note below)	-2.15	2.99	1.55	-7.01	14.99	3.44	-0.55	13.44	17.14	0.45	-2.11	20.44
52	Finance and Insurance	5.48	5.70	0.94	-1.20	13.21	5.81	-0.09	7.88	7.73	0.11	-1.03	9.07
53	Real Estate, Rental	3.23	2.15	0.85	0.17	1.54	2.00	-0.21	-0.24	-1.69	-0.15	-1.05	-0.41
54	Prof., Sci., Tech. Serv.	0.85	3.30	2.16	-4.75	13.30	3.17	-0.32	11.28	12.45	-0.13	-2.48	16.03
56	Admin., support serv.	17.43	2.48	1.32	12.33	-10.02	2.69	-0.70	-11.19	-27.45	0.20	-2.02	-23.52
61	Educational Services	4.88	2.59	3.51	-1.08	2.75	2.30	-1.77	2.21	-2.13	-0.29	-5.29	3.28
62	Health Care, Soc. Ass.	-13.32	-3.37	1.33	-11.24	12.40	-2.98	-1.45	16.15	25.72	0.38	-2.78	27.39
71	Arts, Ent., Rec.	0.80	-0.65	2.62	-1.11	1.25	-0.69	-0.71	2.69	0.45	-0.04	-3.34	3.80
72	Accomm., Food Serv.	1.08	-0.45	0.00	1.62	-2.76	-0.54	0.00	-2.34	-3.83	-0.10	0.00	-3.96
81	Other services	11.63	4.19	4.18	3.10	-1.27	4.03	-1.13	-4.06	-12.89	-0.17	-5.31	-7.17
	Weighted Average	0.94	1.34	1.16	-1.61	4.94	1.47	-0.56	4.10	3.93	0.12	-1.71	5.71
	Weighted Correlation		0.57	0.11	0.94		0.04	-0.31	0.96		0.03	-0.15	0.99

Note: Weighted correlation reports correlation between model prediction and each counterfactual across sectors using the sector's wage bill share in the economy as weights.

account for the aggregate increase in skill observed between 2000 and 2018 in the US, and it is the key driver of the large variation in the evolution in skill observed within sectors in fast- vs. slow-growing firms.

7 Conclusion

Production is non-homothetic, in the sense that the scale of operations drives the worker mix in firms in all sectors of the economy other than transportation and administrative and support services. There is significant variation across sectors in the direction and magnitude of non-homotheticities. In 11 of the 17 non-homothetic sectors, scale leads to an increase in the skill intensity, while in the remaining 6, it leads to an increase in the intensity of un-skilled workers.

To demonstrate these important features of production functions, we have proceeded in three steps. First, we have developed a framework where, other things equal, the scale of operations measured by the efficiency units of labor may affect the optimal skill intensity of firms. Second, we have assembled firm-level datasets that include the relevant observable variables, as well as instrumental variables to identify the causal effect of scale on skill. Third, we have shown that

the causal effect of scale on skill operates through non-homotheticities as the estimates of the skill-elasticity are unaffected by allowing for SBTC to differ across firms at different growth quartiles/deciles.

Non-homotheticities in production have implications for a wide range of significant macroeconomic issues. In this paper, we have applied them to explore the drivers of the evolution of skill intensity. We have shown that variation in scale is a key driver of skill intensity. In particular, changes in scale have been critical to understand the average evolution of skill intensity both at the sector level and the differential evolution of skill between fast- and slow-growing firms.

Failing to recognize the role of scale on the evolution of skill leads to biased estimates of SBTC and the elasticity of substitution between skilled and unskilled workers, and to providing an incorrect account of the drivers of the evolution of skill intensity at the firm, sector and economy-wide levels.

In future work, we intend to explore other implications of production non-homotheticities. For example, our model has implications for the number of efficiency units of labor per worker. This dimension of firm productivity has largely remained unexplored. Variation in efficiency units per worker across firms induces variation in revenue-based measures of productivity (TFPR). In contrast to conventional wisdom, this variation does not reflect market power or other distortions in the economy. Therefore, non-homotheticities also have implications for the measurement of misallocation. Finally, in our study, we have taken the firm-size distribution as given. It would be interesting to explore the joint evolution of scale, productivity and prices in non- homothetic environments.

References

ACEMOGLU, D. AND D. AUTOR (2011): "Skills, Tasks and Technologies: Implications for Employment and Earnings," in *Handbook of Labor Economics*, ed. by O. Ashenfelter and D. Card, Elsevier, vol. 4 of *Handbook of Labor Economics*, chap. 12, 1043–1171.

ADAO, R., M. KOLESÁR, AND E. MORALES (2019): "Shift-Share Designs: Theory and Inference," *Quarterly Journal of Economics*, 134, 1949–2010.

AUTOR, D., D. DORN, AND G. HANSON (2013): "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, 103, 2121–68.

BARANY, Z. AND C. SIEGEL (2018): "Job Polarization and Structural Change," *American Economic Journal: Macroeconomics*, 10, 57–89.

——— (2019): "Job Polarization, Structural Transformation and Biased Technological Change," *Travail et Emploi*, n° 157, 25–44.

——— (2020): "Biased technological change and employment reallocation," *Labour Economics*, 67, S0927537120301342.

BLUNDELL, R. AND J. M. ROBIN (1999): "Estimation in large and disaggregated demand systems: an estimator for conditionally linear systems," *Journal of Applied Econometrics*, 14, 209–232.

BORUSYAK, K., P. HULL, AND X. JARAVEL (2022): "Quasi-experimental shift-share research designs," *The Review of Economic Studies*, 89, 181–213.

COMIN, D. AND M. GERTLER (2006): "Medium-Term Business Cycles," *American Economic Review*, 96, 523–551.

COMIN, D., D. LASHKARI, AND M. MESTIERI (2021): "Structural Change With Long-Run Income and Price Effects," *Econometrica*, 89, 311–374.

COMIN, D. A., A. DANIELI, AND M. MESTIERI (2020): "Income-driven Labor Market Polarization," NBER Working Papers 27455, National Bureau of Economic Research, Inc.

DE LOECKER, J. AND C. SYVERSON (2021): "An industrial organization perspective on productivity," in *Handbook of Industrial Organization, Volume 4*, ed. by K. Ho, A. Hortaçsu, and A. Lizzeri, Elsevier, vol. 4 of *Handbook of Industrial Organization*, 141–223.

HANOCH, G. (1975a): "The Elasticity of Scale and the Shape of Average Costs," *The American Economic Review*, 65, 492–497.

——— (1975b): "Production and Demand Models with Direct or Indirect Implicit Additivity," *Econometrica*, 43, 395–419.

HSIEH, C.-T. AND P. J. KLENOW (2009): "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 124, 1403–1448.

JAEGER, D. A., J. RUIST, AND J. STUHLER (2018): "Shift-Share Instruments and the Impact of Immigration," Tech. Rep. 24285, National Bureau of Economic Research.

JARAVEL, X. (2019): "The Unequal Gains from Product Innovations: Evidence from the U.S. Retail Sector," *Quarterly Journal of Economics*, 134, 715–783.

KATZ, L. AND K. M. MURPHY (1992): "Changes in Relative Wages, 1963–1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, 107, 35–78.

KRUSELL, P., L. E. OHANIAN, J.-V. RÍOS-RULL, AND G. L. VIOLANTE (2000): "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica*, 68, 1029–1053.

LASHKARI, D., A. BAUER, AND J. BOUSSARD (2024): "Information Technology and Returns to Scale," *American Economic Review*, 114, 1769–1815.

MAKSIMOVIC, V. AND G. PHILLIPS (2001): "The Market for Corporate Assets: Who Engages in Mergers and Asset Sales and Are There Efficiency Gains?" *Journal of Finance*, 56, 2019–2065.

——— (2002): "Do Conglomerate Firms Allocate Resources Inefficiently Across Industries? Theory and Evidence," *Journal of Finance*, 57, 721–767.

OBERFIELD, E. AND D. RAVAL (2021): "Micro Data and Macro Technology," *Econometrica*, 89, 703–732.

PERI, G. AND V. YASENOV (2019): "The Labor Market Effects of a Refugee Wave: Synthetic Control Method Meets the Mariel Boatlift," *The Review of Economic Studies*, 89, 181–212.

RESTUCCIA, D. AND R. ROGERSON (2008): "Policy Distortions and Aggregate Productivity with Heterogeneous Plants," *Review of Economic Dynamics*, 11, 707–720.

TROTTNER, F. (2022): "Who gains from scale? Trade and Wage Inequality Within and Between Firms," .

ZEIRA, J. (1998): "Workers, Machines, and Economic Growth," *The Quarterly Journal of Economics*, 113, 1091–1117.

A Appendix

A.1 Proof of proposition 1

Let's define $X = 1 + \frac{\gamma_{Hst}}{(1-\gamma_{Hst})} \left(\frac{W_{Lt}}{W_{Ht}}\right)^{\sigma_s^L} L_{fst}^{(\varepsilon_s^R-1)(1-\sigma_s)} > 1$.

Part a):

$$\frac{\partial L_{fst}/\overline{L}_{fst}}{\partial W_{Lt}/W_{Ht}} = \left(\frac{L_{fst}}{\overline{L}_{fst}}\right)^{\frac{1}{\sigma_s}} [X]^{\frac{1}{\sigma_s}-2} [\sigma_s(X-1)] (1-\gamma_{Hst})^{\frac{1}{\sigma_s}} \left(\frac{W_{Ht}}{W_{Lt}} - 1\right) \tag{30}$$

Therefore $\frac{\partial L_{fst}/\overline{L}_{fst}}{\partial W_{Lt}/W_{Ht}} > 0$ iff $\frac{W_{Ht}}{W_{Lt}} > 1$.

Part b):

$$\frac{\partial L_{fst}/\overline{L}_{fst}}{\partial L_{fst}} = \left(\frac{L_{fst}}{\overline{L}_{fst}}\right)^{\frac{1}{\sigma_s}} \frac{X^{\frac{1}{\sigma_s}-1}}{L_{fst}} \left\{ \left(\frac{L_{fst}}{X}\frac{\partial X}{\partial L_{fst}}\right) (1-\gamma_{Hst})^{\frac{1}{\sigma_s}} \left(\frac{W_{Ht}}{W_{Lt}} - 1\right) \right.$$
$$\left. -(\varepsilon_s^R - 1) \left(\gamma_{Hst} L_{fst}^{(\varepsilon_s^R-1)(1-\sigma_s)}\right)^{\frac{1}{\sigma_s}} (X-1)^{\frac{\sigma_s-1}{\sigma_s}} \right\} \tag{31}$$

where

$$\left(\frac{L_{fst}\partial X}{X\partial L_{fst}}\right) = (\varepsilon_s^R - 1)(1-\sigma_s)\frac{X-1}{X} \tag{32}$$

After some algebra, it follows that

$$\frac{\partial L_{fst}/\overline{L}_{fst}}{\partial L_{fst}} = \left(\frac{L_{fst}}{\overline{L}_{fst}}\right)^{\frac{1}{\sigma_s}} \frac{[X]^{\frac{1}{\sigma_s}-1}(X-1)(1-\gamma_{Hst})^{\frac{1}{\sigma_s}}}{L_{fst}} (1-\varepsilon_s^R)\frac{W_{Ht}}{W_{Lt}} \left[1 - \frac{(1-\sigma_s)}{X}\left(1 - \frac{W_{Lt}}{W_{Ht}}\right)\right] \tag{33}$$

which when $W_{Ht} > W_{Lt}$ is positive iff $\varepsilon_s^R < 1$.

A.2 Cobb-Douglas as Limit to NhCES

TBD

A.3 Additional Tables

Table A1: Joint F-Statistics and Estimated SBTC

		Joint F-statistics		γ_H						
Sector	Sector Title	Rel.Wage	Rel. WB	2000	2003	2006	2009	2012	2015	2018
21	Mining	119.91	50.03	0.01	0.00	0.00	0.00	0.00	0.00	0.01
22	Utilities	470.34	99.95	0.00	0.00	0.00	0.00	0.00	0.00	0.00
23	Construction	1557.25	143.73	0.03	0.03	0.03	0.03	0.04	0.03	0.04
31	Manufacturing	1008.21	243.04	0.95	0.95	0.95	0.95	0.96	0.96	0.95
42	Wholesale Trade	1435.04	207.64	0.20	0.20	0.20	0.22	0.23	0.22	0.23
44	Retail Trade	885.36	465.82	0.19	0.18	0.18	0.19	0.19	0.20	0.19
48	Transportation	228.98	75.18	0.09	0.09	0.08	0.08	0.08	0.08	0.09
49	Warehousing	191.65	22.54	0.53	0.55	0.53	0.56	0.58	0.58	0.56
51	Information (except Tech)	820.74	107.59	0.02	0.02	0.02	0.02	0.02	0.02	0.02
518	Tech (see note below)	163.32	10.89	0.30	0.33	0.32	0.34	0.32	0.32	0.35
52	Finance and Insurance	855.45	473.84	0.31	0.30	0.31	0.31	0.33	0.34	0.36
53	Real Estate, Rental	400.20	79.88	0.31	0.29	0.29	0.30	0.29	0.32	0.33
54	Prof, Sci, Tech. Serv.	4464.47	429.61	0.31	0.33	0.32	0.34	0.34	0.34	0.35
56	Admin, support serv.	728.05	271.80	0.90	0.91	0.91	0.91	0.91	0.92	0.92
61	Educational Services	517.31	745.52	0.63	0.64	0.65	0.64	0.65	0.64	0.66
62	Health Care, Soc. Assis.	14355.40	2922.59	0.03	0.03	0.03	0.03	0.03	0.03	0.03
71	Arts, Ent., Rec.	658.73	338.50	0.23	0.22	0.22	0.23	0.23	0.22	0.22
72	Accomm., Food Serv.	171.90	74.71	0.39	0.37	0.36	0.36	0.38	0.37	0.38
81	Other services	1023.74	107.61	0.82	0.82	0.83	0.83	0.84	0.84	0.85

Note: Mining is short for "Mining, Quarrying, and Oil and Gas Extraction "; Admin. & Support, Waste Mg't is short for "Administrative and Support and Waste Mananagment and Remediation Services."

Table A2: Coefficient of regressing efficiency units per worker on firm scale

		$\rho_{1,s}$	ε_s^R
21	Mining	-0.20277	2.07
22	Utilities	-0.43352	2.57
23	Construction	-0.08674	1.67
31	Manufacturing	-0.23481	2.13
42	Wholesale Trade	-0.14659	2.27
44	Retail Trade	0.028447	0.77
48	Transportation	0.05634	1.02
49	Warehousing	0.240382	0.44
51	Information (exc. Tech)	0.432829	0.34
518	Tech (see note below)	0.41455	0.56
52	Finance and Insurance	0.402079	0.49
53	Real Estate, Rental	0.010558	1.03
54	Prof., Sci., Tech. Serv.	0.4675	0.49
56	Admin., support services	-0.14376	2.18
61	Educ. Serv.	0.165716	0.82
62	Health Care, Soc. Assis.	0.723079	0.12
71	Arts, Ent., Rec.	0.092499	0.79
72	Accomm., Food Serv.	0.047836	0.63
81	Other services	-0.01411	1.24
Weighted Correlation ($\rho_{1,s}$, ε_s^R)			-0.91

Table A3: Model estimates with SBTC specific to each growth decile

	Sector	σ	se_σ	ϵ^R	se_{ϵ^R}	SBTC by growth decile									
						d1	d2	d3	d4	d5	d6	d7	d8	d9	d10
21	Mining	0.15	-	2.01	0.07	2.7%	-62.3%	-58.4%	-56.5%	0.0%	28.4%	23.4%	0.3%	11.3%	60.3%
22	Utilities	0.45	0.12	2.77	0.32	9.3%	-6.7%	-0.3%	10.5%	0.0%	57.5%	47.0%	69.6%	58.7%	112.1%
23	Construction	0.42	0.02	1.74	0.05	19.5%	5.6%	-1.3%	1.4%	0.0%	7.9%	15.3%	24.4%	29.8%	36.8%
31	Manufacturing	2.18	0.02	2.13	0.02	0.5%	0.0%	-0.4%	-0.1%	0.0%	0.5%	0.7%	0.7%	0.8%	1.0%
42	Wholesale Trade	0.88	0.02	1.41	0.19	-3.5%	-12.7%	-13.3%	4.4%	0.0%	21.1%	28.4%	25.2%	29.9%	32.7%
44	Retail Trade	0.59	0.02	0.90	0.05	-12.2%	-21.6%	-16.4%	-5.6%	0.0%	6.1%	10.2%	20.1%	14.7%	5.3%
48	Transportation	0.15	-	1.04	0.04	0.5%	-4.5%	-0.1%	-8.8%	0.0%	-1.3%	1.1%	15.9%	-4.4%	16.4%
49	Warehousing	0.30	0.07	0.51	0.09	28.7%	5.1%	2.6%	18.6%	0.0%	8.0%	-2.0%	15.6%	-10.8%	15.7%
51	Information (ex. Tech)	2.32	0.05	0.36	0.02	22.8%	17.8%	8.2%	20.0%	0.0%	14.4%	28.6%	41.9%	50.6%	49.4%
518	Tech (see note below)	1.85	0.12	0.65	0.10	-4.6%	2.5%	9.1%	7.6%	0.0%	44.7%	15.2%	15.5%	6.5%	19.0%
52	Finance, Insurance	1.36	0.05	0.54	0.11	0.5%	-11.4%	1.5%	14.7%	0.0%	22.3%	26.6%	30.7%	33.5%	29.8%
53	Real Estate, Rental	1.46	0.05	0.92	0.08	-1.2%	-10.5%	-12.9%	-1.4%	0.0%	15.9%	17.9%	24.2%	23.5%	8.1%
54	Prof., Sci., Tech. Serv.	1.67	0.02	0.55	0.03	-1.3%	-12.7%	-9.8%	-3.8%	0.0%	22.6%	25.2%	31.5%	28.2%	25.5%
56	Admin., supp. serv.	1.92	0.06	2.14	0.09	2.5%	2.2%	1.2%	0.2%	0.0%	2.1%	2.1%	2.6%	2.1%	2.8%
61	Educational Serv.	2.01	0.07	0.83	0.03	-0.8%	-6.0%	-8.2%	-0.3%	0.0%	5.7%	7.7%	11.3%	14.6%	13.8%
62	Health C., Soc. Assis.	1.83	0.01	0.15	0.02	-34.9%	-44.4%	-41.5%	-27.6%	0.0%	-8.5%	-5.9%	-2.9%	-6.1%	-1.2%
71	Arts, Ent., Rec.	1.86	0.04	0.82	0.07	-22.4%	-23.5%	-8.7%	-17.3%	0.0%	11.6%	10.7%	12.7%	13.5%	-0.4%
72	Accomm., Food Serv.	1.00	-	0.37	0.08	-6.9%	-12.9%	-15.4%	-7.2%	0.0%	4.0%	10.7%	11.9%	7.5%	5.4%
81	Other services	2.47	0.04	1.23	0.04	2.0%	-0.1%	-0.8%	0.2%	0.0%	5.0%	6.1%	7.3%	7.7%	7.7%

Note: dx denotes the xth decile of the growth rate of WB_{fst}/W_{Lfst}

Table A4: Comparison of estimates in full sample and in restricted sample with less than 10% of employees imputed

		Restricted sample						Full sample					
		Obs.	σ	se_σ	ϵ^R	se_{ϵ^R}	skill el.	Obs	sigma	se_σ	ϵ^R	se_{ϵ^R}	skill el.
21	Mining	1,119	0.15	-	1.39	0.08	0.33	3,822	0.15	-	2.07	0.07	0.91
22	Utilities	4,710	0.26	0.16	2.32	0.24	0.97	10,668	0.41	0.11	2.70	0.28	1.00
23	Construction	23,619	0.45	0.03	1.30	0.08	0.16	71,659	0.42	0.02	1.71	0.05	0.41
31	Manufacturing	65,516	1.94	0.03	2.36	0.04	-1.28	161,252	2.18	0.02	2.13	0.02	-1.33
42	Wholesale Trade	17,011	0.80	0.04	1.15	0.19	0.03	73,780	0.90	0.02	1.74	0.26	0.08
44	Retail Trade	19,054	0.60	0.04	0.72	0.05	-0.11	69,125	0.58	0.02	0.83	0.05	-0.07
48	Transportation	6,149	0.15	-	1.06	0.06	0.05	20,125	0.15	-	1.04	0.04	0.03
49	Warehousing	1,162	0.15	-	0.74	0.06	-0.22	3,885	0.29	0.07	0.49	0.09	-0.36
51	Information (ex. Tech)	6,750	2.16	0.08	0.39	0.04	0.70	18,396	2.32	0.05	0.36	0.02	0.85
518	Tech (see note below)	759	1.68	0.22	0.46	0.25	0.36	2,604	1.86	0.12	0.63	0.10	0.32
52	Finance and Insurance	10,186	1.57	0.12	0.90	0.10	0.06	51,100	1.35	0.04	0.51	0.11	0.17
53	Real Estate, Rental	2,879	1.66	0.09	1.03	0.08	-0.02	11,473	1.47	0.05	0.95	0.08	0.02
54	Prof., Sci., Tech. Serv.	23,747	2.20	0.06	0.88	0.04	0.14	96,824	1.66	0.02	0.52	0.03	0.31
56	Admin., supp. serv.	7,409	1.90	0.09	1.88	0.11	-0.80	25,683	1.92	0.06	2.16	0.09	-1.07
61	Educ. Serv.	8,171	1.76	0.09	0.80	0.06	0.15	19,306	2.01	0.07	0.83	0.03	0.17
62	Health C., Soc. Assist.	59,023	2.46	0.02	0.50	0.02	0.74	170,653	1.82	0.01	0.14	0.02	0.71
71	Arts, Ent., Re.	9,917	1.79	0.05	0.65	0.08	0.27	24,003	1.85	0.04	0.80	0.07	0.17
72	Accom., Food Serv.	2,671	1.00	-	0.67	0.04	-0.33	7,273	1.00	-	0.51	0.05	-0.49
81	Other services	12,443	3.00	0.07	1.28	0.04	-0.56	34,846	2.47	0.04	1.24	0.04	-0.35

Table A5: Variance decomposition of $\ln(\alpha_{Hfst}$ excluding firm random effects)

		$\text{Var}(\ln \alpha_{H,f,t})$	R^2	SBTC	Rel. wages	Scale
21	Mining	0.67	38%	0%	15%	23%
22	Utilities	0.40	29%	3%	-10%	36%
23	Construction	0.46	14%	2%	12%	1%
31	Manufacturing	0.43	15%	0%	5%	10%
42	Wholesale Trade	0.41	4%	1%	1%	2%
44	Retail Trade	0.44	10%	0%	8%	2%
48	Transportation	0.54	34%	0%	34%	0%
49	Warehousing	0.53	34%	0%	37%	-3%
51	Information (exc. Tech)	0.38	30%	1%	-14%	42%
518	Tech (see note below)	0.16	21%	1%	-26%	46%
52	Finance and Insurance	0.23	6%	2%	-6%	10%
53	Real Estate, Rental	0.66	1%	1%	0%	0%
54	Prof., Sci., Tech. Serv.	0.19	7%	1%	-25%	32%
56	Admin., support serv.	0.95	35%	0%	7%	28%
61	Educ. Serv.	0.15	-10%	0%	-28%	18%
62	Health Care, Soc. Assis.	0.63	53%	0%	-17%	70%
71	Arts, Ent., Rec.	0.44	-1%	0%	-2%	1%
72	Accomm., Food Serv.	0.93	0%	0%	0%	0%
81	Other services	0.66	34%	0%	28%	6%
	Weighted Average		19.0%	0.5%	-5.8%	24.2%

Table A6: Estimates of model with quartile-specific SBTC for 3-digit sectors

						Skill-	Growth in $\gamma_{H,s}$		
		σ	se_σ	ϵ^R	se_{ϵ^R}	elast.	Bott. Q.	Midd.	Top Q.
311	Food, beverage, and tobacco	2.30	0.08	1.99	0.04	-1.29	0.5%	0.6%	1.1%
313	Textile, textile product mills	1.91	0.11	2.17	0.14	-1.06	0.1%	-1.1%	0.9%
315	Apparel, and leather	3.86	0.16	1.98	0.06	-2.81	0.0%	0.0%	0.0%
321	Wood product	1.72	0.11	2.77	0.20	-1.27	-1.0%	-1.0%	-1.0%
322	Paper	1.41	0.12	3.30	0.61	-0.93	0.5%	0.0%	3.1%
323	Printing and related support	1.19	0.08	5.06	2.19	-0.77	-0.9%	-0.9%	6.5%
324	Petroleum and coal products	0.31	0.35	1.95	0.33	0.66	-11.6%	25.9%	53.5%
325	Chemical	1.87	0.08	1.29	0.04	-0.25	0.6%	3.7%	8.3%
326	Plastics and rubber products	1.90	0.08	2.38	0.10	-1.24	0.0%	0.6%	1.4%
327	Nonmetallic mineral product	0.64	0.08	0.90	0.17	-0.03	0.6%	-0.1%	3.5%
331	Primary metal	2.13	0.17	1.81	0.07	-0.91	0.1%	0.2%	1.0%
332	Fabricated metal product	1.41	0.05	4.42	0.39	-1.41	-0.6%	0.5%	2.8%
333	Machinery	2.21	0.06	1.92	0.04	-1.12	-0.4%	0.5%	1.3%
334	Computer, electronic prod.	2.33	0.11	0.82	0.07	0.24	13.4%	13.8%	30.6%
335	Electrical equip., appliance	3.05	0.13	1.64	0.04	-1.31	0.2%	0.2%	0.3%
336	Transportation equipment	3.57	0.14	1.45	0.01	-1.15	0.1%	0.2%	0.2%
337	Furniture and related product	1.48	0.08	2.64	0.28	-0.79	3.0%	-0.2%	0.6%
339	Miscellaneous	1.72	0.06	1.99	0.16	-0.71	3.0%	3.0%	6.8%
	Average					-1.00	0.7%	1.5%	4.5%
811	Repair, maintenance	0.76	0.05	0.55	0.22	-0.11	0.8%	-1.2%	21.6%
812	Personal and laundry services	1.64	0.13	1.27	0.08	-0.18	-7.7%	-9.9%	4.1%
8122	Death care services	0.45	0.13	1.13	0.22	0.07	-56.6%	-14.2%	-13.1%
8131	Religious organizations	2.82	0.11	0.85	0.04	0.28	9.3%	13.1%	41.6%
8132	Grantmaking, giving services	3.19	0.21	1.37	0.09	-0.82	0.0%	0.4%	0.7%
8133	Social advocacy organizations	2.59	0.08	0.41	0.07	0.94	-22.8%	-10.2%	19.5%
8134	Civic and social organizations	5.28	0.10	1.79	0.02	-3.36	0.0%	0.0%	0.0%
8139	Business, professional, labor, political organizations	2.55	0.07	0.54	0.04	0.71	12.6%	22.5%	54.7%
	Average					-0.40	-0.3%	5.0%	24.5%

Table A7: Comparison of change in skill Top-Bottom growth quartiles

	Baseline		Quartile-specific SBTC		3-Digit	
Sector	Data	Model	Data	Model	Data	Model
Manufacturing	4.56	-26.42	4.56	-22.31	4.27	-13.42
Other Services	7.99	-12.89	7.99	-4.91	7.86	3.04

Table A8: Counterfactual changes in skill (average across all firms), quartile-specific SBTC

		Model	SBTC	Wages	Scale
21	Mining	1.68	-0.87	-0.18	2.30
22	Utilities	5.60	6.95	-0.35	-1.07
23	Construction	2.34	2.29	-0.47	0.42
31	Manufacturing	2.38	1.86	0.63	0.06
42	Wholesale Trade	3.66	3.49	-0.03	0.14
44	Retail Trade	0.21	0.18	0.19	-0.15
48	Transportation	-0.15	0.14	-0.37	0.08
49	Warehousing	0.04	1.71	-0.71	-1.05
51	Information (exc. Tech)	5.11	4.91	0.81	-0.45
518	Tech (see note below)	8.00	5.27	-0.01	3.65
52	Finance, Insurance	9.26	6.20	0.22	3.05
53	Real Estate, Rental, Leasing	2.51	1.96	0.46	0.06
54	Prof., Sci., Tech. Serv.	7.00	3.86	0.47	3.40
56	Admin., support serv.	2.71	2.99	0.34	-0.72
61	Educational Serv.	4.14	2.51	0.77	0.92
62	Health Care, Social Assist.	-0.15	-3.35	-0.41	3.68
71	Arts, Ent., Rec.	0.68	-0.62	0.64	0.66
72	Accomm., Food Serv.	-0.82	-0.33	0.00	-0.19
81	Other services	4.03	3.02	0.55	0.85
	Average	3.08	1.57	0.17	1.5
	Corr. with model		0.88	0.47	0.25

Note: The model estimates and counterfactuals for manufacturing and other services are at the 3-digit and then aggregated using the share of observations in each 3-digit sector. Average change in skill and correlation between counterfactual and model are computed weighting each sector by its wage bill share in the economy-wide wage bill.

Table A9: Counterfactual of differential changes in skill in fast-vs. slow growing firms, quartile-specific SBTC

		Model	SBTC	Wages	Scale
21	Mining	27.13	11.28	0.87	14.89
22	Utilities	21.01	16.82	1.07	2.84
23	Construction	11.92	3.14	0.70	7.17
31	Manufacturing	-13.42	3.65	-1.53	-14.51
42	Wholesale Trade	13.72	11.78	0.33	1.49
44	Retail Trade	4.70	5.01	1.10	-1.39
48	Transportation	2.80	1.16	1.00	0.66
49	Warehousing	-13.44	-2.31	0.77	-10.83
51	Information (exc. Tech)	36.16	5.99	-4.33	34.72
518	Tech (see note below)	19.97	6.23	-1.97	16.64
52	Finance,Insurance	22.43	14.05	-1.04	10.24
53	Real Estate, Rental, Leasing	5.79	6.19	-1.08	0.60
54	Prof., Sci., Tech. Serv.	23.81	10.82	-2.44	17.55
56	Admin., support serv.	-27.62	0.00	-1.96	-22.96
61	Educational Serv.	6.85	9.02	-5.25	3.33
62	Health Care, Social Assist	36.28	6.62	-2.85	32.10
71	Arts, Ent., Rec.	7.94	7.35	-3.49	3.79
72	Accomm., Food Serv.	-1.15	4.30	0.00	-5.64
81	Other services	3.04	6.26	-5.14	3.44
	Average	13.41	7.25	-1.68	8.31
	Corr. with model		0.52	-0.28	0.97

Note: The model estimates and counterfactuals for manufacturing and other services are at the 3-digit and then aggregated using the share of observations in each 3-digit sector. Average change in skill and correlation between counterfactual and model are computed weighting each sector by its wage bill share in the economy-wide wage bill.