Web Appendix for Tasks and Technology in the United States 1880-2000 (Not for Publication)*

Guy Michaels
London School of Economics and CEP[†]
Ferdinand Rauch
University of Oxford and CEP[‡]
Stephen J. Redding
Princeton University and CEP[§]

February 16, 2016

Contents

*We are grateful to the Centre for Economic Performance and Princeton University for research support. The usual disclaimer applies.

[†]CEP, LSE, Houghton Street, WC2A 2AE, London, UK. Email: g.michaels@lse.ac.uk. Tel: +44(0)20-7852-3518.

[‡]Department of Economics, University of Oxford, Manor Road, Oxford, OX1 3UQ. Email: ferdinand.rauch@economics.ox.ac.uk. Tel: +44(0)77214-88037.

[§]Fisher Hall, Princeton, NJ 08540. Email: reddings@princeton.edu. Tel: +1(609) 258-4016.

A1 Introduction

This technical appendix contains additional supplementary material for the paper. Section A2 contains a more detailed analysis of the model, including proofs of propositions. Section A3 develops the extension to multiple worker types discussed in Section 2 of the paper. Section A4 presents the extension to multiple locations referred to in Section 2 of the paper. Section A5 provides some further detail on the data sources and definitions. Section A6 reports additional empirical results for Section 5 of the paper. Section A7 contains further empirical results for Section 6 of the paper. Section A8 presents supplementary empirical results for Section 7 of the paper. Section A9 reports additional robustness tests and empirical specifications discussed in the paper.

A2 Model

In this section, we develop in further detail the theoretical model outlined in the paper. We present the complete technical derivations of all the expressions and results reported in the paper. In the interests of clarity and to ensure that this section of the web appendix is self-contained, we reproduce some material from the paper, but also include the intermediate steps for the derivation of expressions.

To guide our empirical analysis, we develop a simple Roy model, in which workers endogenously sort across occupations and sectors based on their comparative advantage.[1] We begin by developing a baseline version of the model in which workers are *ex ante* identical and the economy consists of a single location (e.g. the U.S.). We next demonstrate the robustness of our results to two extensions, one to allow for multiple worker types that differ *ex ante* in observable characteristics (e.g. gender, age and general schooling), and the other to allow for multiple locations (e.g. urban and rural areas), as in our data.

[1] We extend the version of the Roy model in Hsieh, Hurst, Jones, and Klenow (2013) to incorporate multiple sectors and locations, as observed in our data. The classic treatments of the Roy model are Roy (1951) and Heckman and Honore (1990). See also Lagakos and Waugh (2013), Ahlfeldt, Redding, Sturm, and Wolf (2015) and Burstein, Morales, and Vogel (2015).

We use the model to show how changes in technology affect employment shares and average wages in a setting where workers self-select across many sectors and occupations based on idiosyncratic realizations for ability. We highlight two mechanisms through which technology affects employment shares and wages: the average effectiveness of workers in performing tasks and/or the rate of return to human capital accumulation. Under our assumption of a Fréchet distribution for idiosyncratic ability, changes in the average rate of return to human capital accumulation only affect average wages, whereas changes in average task effectiveness affect both employment shares and average wages. Guided by these predictions, our empirical work examines the extent to which employment shares and average wages have changed systematically towards occupations performing certain types of tasks (measured using both numerical scores and our new methodology); the extent to which these changes are related to direct measures of new technologies (e.g. information and communication technologies); and the extent to which these changes differ between urban and rural areas.

The economy consists of a continuum of people ($\bar{L}$) who can choose to work in O possible occupations. Human capital for each occupation depends on raw worker ability and investments in human capital accumulation for that occupation. People choose an occupation based on the wage and cost of investing in occupational human capital. Occupational human capital is used to produce final goods in S sectors. We allow some occupations (e.g. managers) to be employed in most sectors, while other occupations (e.g. lathe operators) may be employed in only a few sectors.

A2.1 Preferences and Technology

A person i with consumption C_i and leisure time $1 - \ell_i$ obtains utility:

$$U_i = C_i^{\beta} (1 - \ell_i), \qquad \beta > 0, \tag{A1}$$

where C_i is a consumption index; ℓ_i represents investments in human capital accumulation; and β parameterizes the tradeoff between consumption and the accumulation of human capital. The consumption index (C_i) is itself a Cobb-Douglas function of consumption of tradeable goods (C_{Mi}) and a non-tradeable good (C_{Ni}) that we interpret as housing:[2]

$$C_i = \left(\frac{C_{Mi}}{\alpha} \right)^{\alpha} \left(\frac{C_{Ni}}{1 - \alpha} \right)^{1-\alpha}, \qquad 0 < \alpha < 1, \tag{A2}$$

where housing is assumed to be in inelastic supply $\bar{N}$ and the presence of this non-traded good ensures a non-degenerate distribution of economic activity in the multi-region version of the model below.

The tradeables consumption index (C_{Mi}) is a constant elasticity of substitution (CES) function of consumption of a number of sectors (C_{Mis}) indexed by $s \in \{0, \dots, S\}$:

$$C_{Mi} = \left[\sum_{s=1}^{S} (\zeta_s C_{Mis})^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \qquad \sum_{s=1}^{S} \zeta_s = 1, \tag{A3}$$

[2] For empirical evidence in support of the constant housing expenditure share implied by this Cobb-Douglas functional form, see Davis and Ortalo-Magné (2011).

where ζ_s controls the strength of relative preferences for sector s and σ is the elasticity of substitution between sectors. Output in each tradeable sector (Y_{Ms}) is a constant elasticity of substitution (CES) function of the human capital of workers from each occupation within that sector (H_{so}):

$$Y_{Ms} = \left[\sum_{o=1}^{O} \left(\xi_{so} H_{so} \right)^{\frac{\kappa-1}{\kappa}} \right]^{\frac{\kappa}{\kappa-1}}, \qquad \sum_{o=1}^{O} \xi_{so} = 1 \; \forall \; s, \tag{A4}$$

where ξ_{so} controls the relative productivity of occupation o in sector s; this occupation o is not employed in sector s if $\xi_{so} = 0$; κ is the elasticity of substitution between occupations within sectors; and goods market clearing requires that output of each good equals the sum of all individuals' consumption of that good: $Y_{Ms} = \sum_i C_{Mis}$.[3]

Workers choose an occupation and acquire human capital for that occupation, such that each occupation corresponds to a separate labor market. Workers within each occupation o are mobile across sectors, which implies that the wage per effective unit of labor within that occupation (w_o) is the same across sectors. Each worker i's choice of sector s within occupation o is determined by their idiosyncratic realizations for effective units of labor (ability z_{iso}) for each sector and occupation. Each worker i's choice of occupation o depends on these realizations for idiosyncratic ability, the wage per effective unit of labor for each occupation, and the rate of return to human capital investments for each occupation.

Each person i works one unit of time in her chosen occupation $o \in \{1, \ldots, O\}$. Another unit of time is divided between leisure $(1 - \ell_{io})$ and human capital accumulation (ℓ_{io}). The production function for human capital in occupation o is:

$$h_{io} \left(\ell_{io} \right) = \bar{h}_o \ell_{io}^{\phi_o}, \tag{A5}$$

where the parameter $\bar{h}_o > 0$ captures the productivity of human capital investments; the parameter $\phi_o > 0$ determines the rate of return to human capital accumulation; and both parameters can differ across occupations o.

Human capital for each sector and occupation (H_{so}) equals the fraction of agents who choose that sector and occupation (λ_{so}) times average human capital conditional on choosing that sector and occupation times the measure of agents in the economy ($\bar{L}$):

$$H_{so} = \lambda_{so} \, \mathbb{E} \left[h_o z \mid \text{Person chooses } s \text{ and } o \right] \bar{L}. \tag{A6}$$

where average human capital depends on both human capital accumulation (h_o) and ability (z).

Each person's income depends on the wage per effective unit of labor for her chosen occupation (w_o), her accumulated human capital for that occupation (h_{io}) and her idiosyncratic ability (z_{iso}) for her chosen sector and occupation:

$$\Omega_i = w_o h_{io} z_{iso} = w_o \bar{h}_o \ell_{io}^{\phi_o} z_{iso}. \tag{A7}$$

[3] We assume for simplicity that labor is the sole factor of production, but the analysis can be extended to incorporate other factors of production such as capital or land. We also assume for simplicity that κ takes the same value across sectors, but it is straightforward to allow this elasticity to differ across sectors.

The timing of decisions is as follows. First, each person i observes her realizations of idiosyncratic ability (z_{iso}) and chooses a sector s and occupation o, taking occupational wages (w_o) as given. Second, she chooses her optimal human capital investment in her chosen occupation (ℓ_{io}), given the trade-off between goods consumption and human capital accumulation in utility (A1) and the technology for accumulating human capital (A5). Third, she makes her optimal choices for overall goods consumption (C_{Mi}), consumption of housing (C_{Ni}), and goods consumption for each sector (C_{Mis}), given observed prices (P_{Ms}, P_N) and her income (Ω_i) in her chosen sector and occupation (A7), as determined by wages, human capital investments and idiosyncratic ability.

An equilibrium in this economy is a set of allocations of consumption, production, human capital investments and choices of sector and occupation $\{C_i, C_{Mi}, C_{Ni}, C_{Msi}, Y_{Msi}, \ell_{io}, \lambda_{so}\}$ and a set of prices $\{P_M, P_N, P_{Ms}, w_o\}$, such that individuals choose consumption, human capital investments, sector and occupation to maximize utility; firms choose inputs of human capital to maximize profits; zero profits are made if a good is produced; and the markets for goods, labor and housing clear. We use the timing of decisions and structure of the model to solve for equilibrium recursively. First, we characterize equilibrium consumption and production as a function of human capital investments and choice of sector and occupation. Second, we determine optimal human capital investments as a function of choice of sector and occupation. Third, we solve for the optimal choice of sector and occupation.

A2.2 Consumption Decisions

Given an individual's human capital investments and her choice of sector and occupation, the characterization of consumption decisions is straightforward. The Cobb-Douglas functional form (A2) implies that each person allocates constant shares of income to consumption of goods and housing: $C_{Mi} = \alpha\Omega_i/P_M$ and $C_{Ni} = (1-\alpha)\Omega_i/P_N$. Using these results in (A1), the utility function can be written in terms of income, the prices of goods consumption and housing, and investment in human capital accumulation:

$$U_i = \left(\frac{w_o \bar{h}_o \ell_{io}^{\phi_o} z_{iso}}{P_M^{\alpha} P_N^{1-\alpha}} \right)^{\beta} (1 - \ell_{io}) \,. \tag{A8}$$

Using the CES functional form of goods consumption (A3), the share of goods expenditure allocated to each sector (μ_{Ms}) depends on prices (P_{Ms}) and preference weights (ζ_s):

$$\mu_{Ms} = \frac{P_{Ms} C_{Ms}}{\sum_{k=1}^{S} P_{Mk} C_{Mk}} = \frac{(P_{Ms}/\zeta_s)^{1-\sigma}}{\sum_{k=1}^{S} (P_{Mk}/\zeta_k)^{1-\sigma}}, \tag{A9}$$

and the overall price index (P_M) dual to the goods consumption index (A3) can be written in terms of prices (P_{Ms}) and preference weights (ζ_s) for each sector:

$$P_M = \left[\sum_{s=1}^{S} \left(\frac{P_{Ms}}{\zeta_s} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} . \tag{A10}$$

Using the CES production technology (A4), the share of each occupation in the wage bill in each sector depends on occupational wages (w_o) and productivities (ξ_{so}):

$$\frac{w_o H_{so}}{\sum_{m=1}^{O} w_m H_{sm}} = \frac{(w_o/\xi_{so})^{1-\kappa}}{\sum_{m=1}^{O} (w_m/\xi_{sm})^{1-\kappa}}. \tag{A11}$$

Perfect competition implies that prices in each sector (P_{Ms}) equal unit costs, which can be expressed as the following function of occupational wages (w_o) and productivities (ξ_{so}):

$$P_{Ms} = \left[\sum_{o=1}^{O} \left(\frac{w_o}{\xi_{so}}\right)^{1-\kappa}\right]^{\frac{1}{1-\kappa}}. \tag{A12}$$

The occupational wage (w_o) is determined by the requirement that the total income of all workers within the occupation is equal to total payments to workers in that occupation:

$$w_o \sum_{s=1}^{S} H_{so} = \sum_{s=1}^{S} \frac{(w_o/\xi_{so})^{1-\kappa}}{\sum_{m=1}^{O} (w_m/\xi_{sm})^{1-\kappa}} \mu_{Ms} E. \tag{A13}$$

Finally, the price of housing is determined by housing market clearing:

$$P_N = \frac{(1-\alpha)\Omega}{\bar{N}}, \tag{A14}$$

where recall that $\bar{N}$ is the inelastic supply of housing and aggregate income (Ω) is the sum of each person's income (Ω_i). The larger the measure of people in the economy, and the higher their income relative to the supply of housing, the higher the price of housing.

A2.3 Human Capital Investments

Given a choice of sector s and occupation o, wages (w_o), realizations of idiosyncratic ability (z_{iso}), and prices of tradeable (P_M) and non-tradeable (P_N) goods, each individual i chooses her human capital investment (ℓ_{io}) to maximize her utility (A8):

$$\max_{\ell_{io}} \left\{ \left(\frac{w_o \bar{h}_o \ell_{io}^{\phi_o} z_{iso}}{P_M^{\alpha} P_N^{1-\alpha}} \right)^{\beta} (1-\ell_{io}) \right\}. \tag{A15}$$

The first-order condition to this problem yields the equilibrium human capital investment:

$$\ell_o^* = \frac{1}{1+\frac{1}{\beta\phi_o}}, \tag{A16}$$

which only varies across occupations o and not across individuals i within occupations. Hence, from now onwards, we suppress the individual subscript i, unless otherwise indicated.

Equilibrium human capital investments depend solely on β (the tradeoff between consumption and the accumulation of human capital) and ϕ_o (the productivity of human capital accumulation in the worker's chosen occupation). Other forces do not affect these investments, because they have the same effect on the return and opportunity cost to human capital accumulation. The expression (A16) highlights that a first key mechanism through which technological change can affect the economy is that it can change the productivity of human capital investments (ϕ_o) in some occupations relative to others. For example, computers may increase the ease of acquiring analytical and technical skills used in engineering occupations relative to manual skills used in laboring occupations.

A2.4 Sector and Occupation Choice

Given occupational wages (w_o), realizations of idiosyncratic ability in each sector and occupation (z_{iso}), and prices of tradeable (P_M) and non-tradeable (P_N) goods, each individual chooses her sector and occupation to maximize her utility. We model individual ability following McFadden (1974) and Eaton and Kortum (2002). Each individual i draws ability for each sector s and occupation o (z_{iso}) from an independent Fréchet distribution:

$$F_{so}(z) = e^{-T_{so}z^{-\theta}}, \qquad \theta > 1. \tag{A17}$$

The Fréchet scale parameter T_{so} determines average effective units of labor for workers in sector s and occupation o, which we refer to as the average effectiveness of workers in performing tasks in that sector and occupation. The Fréchet shape parameter θ determines the dispersion of effective units of labor across sectors and occupations. A reduction in θ corresponds to an increased dispersion of effective units of labor and greater scope for worker specialization according to comparative advantage across sectors and occupations.[4] In this specification (A17), another mechanism through which technological change can affect the economy is that it can raise the effectiveness of workers in performing tasks in some sectors and occupations (T_{so}) relative to others. For example, on the one hand, computers may complement workers in performing design and simulation tasks in engineering occupations. On the other hand, computers may substitute for workers in performing routine calculations in clerical occupations. In our empirical work, we use the structure of the model to estimate the extent to which new technologies complement or substitute for different occupations by changing the average effectiveness of workers in performing different tasks (e.g. the formation of ideas versus the manipulation of the physical world).[5]

From utility (A8), the following transformation of utility is linear in worker ability:

$$v = U^{1/\beta} = \bar{w}_o z, \qquad \bar{w}_o = \frac{w_o (1-\ell_o)^{1/\beta} \ell_o^{\phi_o} \bar{h}_o}{P_M^{\alpha} P_N^{1-\alpha}}. \tag{A18}$$

Using this monotonic relationship between utility and worker ability, the distribution of utility across

[4] Although we assume that ability is drawn independently for each sector and occupation, the parameter T_{so} induces a correlation in ability among workers within the same sector and occupation. While we focus on the independent Fréchet distribution for simplicity, it is straightforward to instead consider the multivariate Fréchet distribution, which allows for correlation in the ability draws of individual workers across sectors and occupations.

[5] Our theoretical framework models technological change as determining the relative effectiveness of workers in performing tasks in different sectors and occupations. Alternatively, technological change could be modeled as embodied in physical capital and machines. In both cases, new technologies can either complement or substitute for workers in particular occupations. Our approach enables us to tractably model the effects of technological change in a setting with many sectors and occupations in the context of a standard Roy model.

workers within each sector s and occupation o inherits a Fréchet distribution:

$$F_{so}(v) = e^{-\Phi_{so} v^{-\theta}}, \qquad \Phi_{so} = T_{so} \bar{w}_o^{\theta}. \tag{A19}$$

Each worker chooses their occupation and sector to maximize their utility. Note that the maximum of Fréchet distributed random variables also has a Fréchet distribution. Therefore the distribution of utility across all sectors and occupations is given by:

$$F(v) = e^{-\Phi v^{-\theta}}, \qquad \Phi = \sum_{s=1}^{S} \sum_{o=1}^{O} T_{so} \bar{w}_o^{\theta}, \tag{A20}$$

where the Fréchet functional form implies that the distribution of utility conditional on choosing a sector and occupation is the same for each sector and occupation pair and equal to the distribution of utility across all sectors and occupations (A20).

Aggregating optimal choices of occupation and sector across people, we arrive at our first key result for equilibrium worker sorting across sectors and occupations.

Proposition 1 ***(Sector and Occupation Choice)*** *Let λ_{so} denote the fraction of people who choose to work in sector s and occupation o. Let λ_o denote the fraction of people who choose to work in occupation o. Aggregating across people, the model yields the following sufficient statistics for the fractions of people choosing to work in each sector and occupation (Ψ_{so}) and in each occupation (Ψ_o):*

$$\lambda_{so} = \frac{\Psi_{so}}{\Psi}, \qquad \lambda_o = \frac{\Psi_o}{\Psi}, \tag{A21}$$

$$\Psi_o = \sum_{s=1}^{S} \Psi_{so}, \qquad \Psi = \sum_{s=1}^{S} \sum_{o=1}^{O} \Psi_{so}, \qquad \Psi_{so} = T_{so} w_o^{\theta} (1 - \ell_o)^{\theta/\beta} \ell_o^{\theta \phi_o} \bar{h}_o^{\theta}.$$

Proof. The probability that a worker chooses sector s and occupation o is:

$$\begin{aligned}
\lambda_{so} =& \Pr\left[v_{so} \geq \max\{v_{km}\}; \forall k, m\right], \\
=& \int_0^{\infty} \prod_{k \neq s} F_{ko}(v) \left[\prod_{k=1}^{S} \prod_{m \neq o} F_{km}(v)\right] f_{so}(v) dv, \\
=& \int_0^{\infty} \prod_{k \neq s} e^{-\Phi_{ko} v^{-\theta}} \left[\prod_{k=1}^{S} \prod_{m \neq o} e^{-\Phi_{km} v^{-\theta}}\right] \theta \Phi_{so} v^{-(\theta+1)} e^{-\Phi_{so} v^{-\theta}} dv, \\
=& \int_0^{\infty} \prod_{k=1}^{S} \prod_{m=1}^{O} \theta \Phi_{so} v^{-(\theta+1)} e^{-\Phi_{km} v^{-\theta}} dv, \\
=& \int_0^{\infty} \theta \Phi_{so} v^{-(\theta+1)} e^{-\Phi} v^{-\theta} dv.
\end{aligned}$$

Note that:

$$\frac{d}{dv}\left[-\frac{1}{\Phi} e^{-\Phi v^{-\theta}}\right] = \theta v^{-(\theta+1)} e^{-\Phi}.$$

Therefore:

$$\lambda_{so} = \left[-\frac{\Phi_{so}}{\Phi} e^{-\Phi v^{-\theta}} \right]_0^{\infty},$$

which becomes:

$$\lambda_{so} = \frac{\Phi_{so}}{\Phi} = \frac{\Psi_{so}}{\Psi},$$

where

$$\frac{\Phi_{so}}{\Phi} = \frac{\frac{T_{so} w_o^{\theta} (1-\ell_o)^{\theta/\beta} \ell_o^{\theta\phi_o} \bar{h}_o^{\theta}}{P_M^{\alpha\theta} P_N^{(1-\alpha)\theta}}}{\sum_{k=1}^{S} \sum_{m=1}^{O} \frac{T_{km} w_m^{\theta} (1-\ell_m)^{\theta/\beta} \ell_m^{\theta\phi_m} \bar{h}_m^{\theta}}{P_M^{\alpha\theta} P_N^{(1-\alpha)\theta}}},$$

which simplifies to:

$$\frac{\Psi_{so}}{\Psi} = \frac{T_{so} w_o^{\theta} \left(1-\ell_o\right)^{\theta/\beta} \ell_o^{\theta\phi_o} \bar{h}_o^{\theta}}{\sum_{k=1}^{S} \sum_{m=1}^{O} T_{km} w_m^{\theta} \left(1-\ell_m\right)^{\theta/\beta} \ell_m^{\theta\phi_m} \bar{h}_m^{\theta}}.$$

Summing across sectors s, we obtain the probability that a worker chooses occupation o:

$$\lambda_o = \sum_{s=1}^{S} \lambda_{so} = \frac{\Psi_o}{\Psi}, \qquad \Psi_o = \sum_{s=1}^{S} \Psi_{so} = \sum_{s=1}^{S} T_{so} w_o^{\theta} \left(1-\ell_o\right)^{\theta/\beta} \ell_o^{\theta\phi_o} \bar{h}_o^{\theta}.$$

■

Therefore the fraction of workers choosing to work in each sector and occupation depends on average effective units of labor (as determined by T_{so}), the wage in each occupation (w_o), and equilibrium human capital investments (as determined by ℓ_o, which in turn depends solely on β and ϕ_o). Note that the terms in the tradeables consumption price (P_M) and the non-tradeables price (P_N) in Φ_{so} in (A18)-(A19) have cancelled from the choice probabilities λ_{so} in (A21), because they are common across sectors and occupations, and hence do not affect the choice of sector and occupation.

Thus the model suggests that employment shares are one of the key endogenous variables of interest in our empirical work. The model highlights the role of new technologies in influencing these employment shares through changing the effectiveness of workers in performing the tasks in different sectors and occupations. Changes in technology have both direct effects on employment shares (though T_{so}) and indirect effects (through general equilibrium effects via occupation wages w_o). Totally differentiating the sector-occupation choice probabilities (A21), holding constant the rate of return to human capital investments (ϕ_o and hence ℓ_o), we have:

$$\frac{d\lambda_{so}}{\lambda_{so}} = \frac{dT_{so}}{T_{so}} - \sum_{k=1}^{S} \sum_{m=1}^{O} \frac{dT_{km}}{T_{km}} \lambda_{km} + \theta \frac{dw_o}{w_o} - \sum_{k=1}^{S} \sum_{m=1}^{O} \theta \frac{dw_m}{w_m} \lambda_{km}. \tag{A22}$$

Evaluating these total derivatives holding occupational wages constant at their values in the initial equilibrium (setting $dw_o/w_o = 0$ for all occupations o), the direct effect of an increase in the effectiveness of workers in performing the tasks in a sector and occupation (T_{so}) is to raise the share of employment (λ_{so}) in that sector and occupation (since $0 < \lambda_{so} < 1$) and reduce the share of employment in all other sectors and occupations. Therefore, to the extent that new technologies complement workers in performing tasks in a sector and occupation (higher T_{so}), we would expect them to increase employment shares in that sector and occupation (higher λ_{so}), other things equal. In contrast, to the extent that new technologies substitute for workers in performing tasks in a sector and occupation (lower T_{so}), we would expect them to decrease employment shares in that sector and occupation (lower λ_{so}), other things equal. We provide evidence in the paper that the impact of new technologies on the employment shares of different occupations is related to the production tasks performed by workers within those occupations.

Our second key result for equilibrium worker sorting is for average earnings in each occupation.

Proposition 2 ***(Occupational Average Earnings)*** *The model's sufficient statistic for average earnings in occupation o ($\overline{wage}_o$), including both human capital and ability, is:*

$$\overline{wage}_o = \mathbb{E}\left[w_o h_o z\right] = \gamma \left(1-\ell_o\right)^{-1/\beta} \left(P_M^{\alpha} P_N^{1-\alpha}\right) \Phi^{1/\theta}, \tag{A23}$$

where $\gamma = \Gamma\left(\frac{\theta-1}{\theta}\right)$ and $\Gamma\left(\cdot\right)$ is the Gamma function.

Proof. Average earnings in a person's chosen occupation depends on the wage per effective unit of labor, her human capital accumulation and her average ability conditional on choosing that occupation:

$$\overline{wage}_o = \mathbb{E}\left[w_o h_o z \mid \text{chooses } o\right]. \tag{A24}$$

To derive the distribution of ability conditional on choosing an occupation, note that the distribution of utility conditional on choosing an occupation is the same as the distribution of utility across all sectors and occupations (A20), reproduced below:

$$F\left(v\right) = e^{-\Phi v^{-\theta}}, \qquad \Phi = \sum_{k=1}^{S} \sum_{m=1}^{O} T_{km} \bar{w}_m^{\theta}.$$

Using the monotonic relationship between utility and ability:

$$v = \bar{w}_o z,$$

we obtain the distribution of ability conditional on choosing an occupation:

$$F(z) = e^{-\Phi^* z^{-\theta}}, \qquad \Phi^* = \sum_{k=1}^{S} \sum_{m=1}^{O} T_{km} \left(\bar{w}_m / \bar{w}_o\right)^{\theta}.$$

Therefore expected ability in a worker's chosen occupation is:

$$\mathbb{E}\left[z \mid \text{chooses } o\right] = \int_0^{\infty} \theta \Phi^* z^{-\theta} e^{-\Phi^* z^{\theta}} dz.$$

Defining the following change of variables:

$$y = \Phi^* z^{-\theta}, \qquad dy = \theta \Phi^* z^{-(\theta+1)} dz,$$

expected ability can be re-written as:

$$\mathbb{E}\left[z \mid \text{chooses } o\right] = \int_0^{\infty} y^{-1/\theta} \left(\Phi^*\right)^{1/\theta} e^{-y} dy,$$

which yields:

$$\mathbb{E}\left[z \mid \text{chooses } o\right] = \gamma \left(\Phi^*\right)^{1/\theta}, \qquad \gamma = \Gamma\left(\frac{\theta-1}{\theta}\right), \tag{A25}$$

where $\Gamma(\cdot)$ is the gamma function. Now note that:

$$\lambda_o = \frac{\sum_k T_{ko}\bar{w}_o^\theta}{\sum_{k=1}^S \sum_{m=1}^O T_{km}\bar{w}_m^\theta}, \tag{A26}$$

which can be written as:

$$\lambda_o = \frac{\sum_k T_{ko}}{\sum_{k=1}^S \sum_{m=1}^O T_{km}\left(\bar{w}_m/w_o\right)^\theta} = \frac{\sum_k T_{ko}}{\Phi^*}.$$

Therefore expected ability in a worker's chosen occupation (A25) can be re-expressed as:

$$\mathbb{E}\left[z \mid \text{chooses } o\right] = \gamma \left(\frac{\sum_{k=1}^S T_{ko}}{\lambda_o}\right)^{1/\theta}, \qquad \gamma = \Gamma\left(\frac{\theta-1}{\theta}\right). \tag{A27}$$

It follows that average human capital in a worker's chosen occupation is:

$$\mathbb{E}\left[h_o z \mid \text{chooses } o\right] = \gamma \bar{h}_o \ell_o^{\phi_o} \left(\frac{\sum_{k=1}^S T_{ko}}{\lambda_o}\right)^{1/\theta}. \tag{A28}$$

Using the occupational choice probability (A26), average human capital in a worker's chosen occupation can be re-written as:

$$\mathbb{E}\left[h_o z \mid \text{chooses } o\right] = \gamma \bar{h}_o \ell_o^{\phi_o} \frac{1}{\bar{w}_o} \left(\sum_{k=1}^S \sum_{m=1}^O T_{km}\bar{w}_m^\theta\right)^{1/\theta}. \tag{A29}$$

Using the definition of $\bar{w}_o$ in (A18), it follows that average income in a worker's chosen occupation is:

$$\mathbb{E}\left[w_o h_o z \mid \text{chooses } o\right] = \gamma \left(1-\ell_o\right)^{-1/\beta} P_M^\alpha P_N^{1-\alpha} \left(\sum_{k=1}^S \sum_{m=1}^O T_{km}\bar{w}_m^\theta\right)^{1/\theta}. \tag{A30}$$

Therefore relative average income only varies across occupations because of variation in human capital investments ℓ_o:

$$\frac{\mathbb{E}\left[w_o h_o z \mid \text{chooses } o\right]}{\mathbb{E}\left[w_m h_m z \mid \text{chooses } m\right]} = \left(\frac{1-\ell_o}{1-\ell_m}\right)^{-1/\beta}. \tag{A31}$$

■

Therefore differences in average earnings ($\overline{wage}_o$) across occupations o are explained in the model by differences in human capital investments ($(1-\ell_o)^{-1/\beta}$). Occupations in which human capital investments are more productive (higher ϕ_o) have higher human capital investments (ℓ_o) and higher average earnings ($\overline{wage}_o$). In contrast, average earnings are no higher in occupations that have higher average effective units of labor (higher T_{so}) or higher wages per effective unit of labor (higher w_o). The reason is a selection effect. On the one hand, higher T_{so} and w_o directly *increase* average wages for a given fraction of workers choosing to enter an occupation. On the other hand, higher T_{so} and w_o induce a higher fraction of workers to choose an occupation, which indirectly *reduces* average wages through a composition effect of a higher fraction of workers with lower draws for effective units of labor. With a Fréchet distribution for worker ability, these two effects exactly offset one another, leaving average earnings unchanged. Although this exact offset is a feature of the Fréchet distribution, more generally, these two effects work in different directions and dampen the impact of T_{so} and w_o on average earnings ($\overline{wage}_o$).

Thus the model highlights average earnings as the second key endogenous variable of interest in our empirical work. New technologies affect average earnings through changing the return to human capital accumulation in different occupations. Totally differentiating average earnings (A23), changes in the relative average earnings ($\omega_{om} = \overline{wage}_o / \overline{wage}_m$) of two occupations o and m depend solely on changes in human capital investments:

$$\frac{d\omega_{om}}{\omega_{om}} = \frac{1}{\beta}\left[\frac{\ell_o}{1-\ell_o}\frac{d\ell_o}{\ell_o} - \frac{\ell_m}{1-\ell_m}\frac{d\ell_m}{\ell_m}\right], \tag{A32}$$

where changes in these human capital investments depend solely on changes in the rate of return to these investments (ϕ_o):

$$\frac{d\ell_o}{\ell_o} = \frac{1}{1+\beta\phi_o}\frac{d\phi_o}{\phi_o}, \qquad \frac{d\ell_m}{\ell_m} = \frac{1}{1+\beta\phi_m}\frac{d\phi_m}{\phi_m}. \tag{A33}$$

Therefore, to the extent that new technologies are complementary to human capital investments within an occupation (higher ϕ_o), we would expect them to increase occupation average earnings (higher $\overline{wage}_o$). In contrast, to the extent that new technologies substitute for human capital investments within an occupation (lower ϕ_o), we would expect them to decrease occupation average earnings (lower $\overline{wage}_o$).

A2.5 Quantification

We now show how observed values of the two key endogenous variables in the model, employment shares and average earnings, can be used to solve for unobserved values of the rate of return to human capital accumulation and an adjusted measure of the average effectiveness of workers of performing tasks in each sector and occupation. We first assume central values for the model's parameters from the existing empirical literature. We follow Hsieh, Hurst, Jones, and Klenow (2013) in assuming a value for the Fréchet shape parameter determining worker comparative advantage of $\theta = 3.44$ and a value for the parameter governing the tradeoff between consumption and human capital accumulation of $\beta = 0.693$. Using these assumed parameters and the expressions for equilibrium human capital investments (A16) and average earnings (A23), we can recover the return to human capital accumulation from observed average earnings in each occupation relative to the geometric mean of average earnings across occupations:

$$\frac{\overline{wage}_o}{\prod_{o=1}^{O}\left[\overline{wage}_o\right]^{\frac{1}{O}}} = \frac{(1-\ell_o)^{-1/\beta}}{\prod_{o=1}^{O}\left[(1-\ell_o)^{-1/\beta}\right]^{\frac{1}{O}}} = \frac{(1+\beta\phi_o)^{-1/\beta}}{\prod_{o=1}^{O}\left[(1+\beta\phi_o)^{-1/\beta}\right]^{\frac{1}{O}}}. \tag{A34}$$

Given observed occupation average earnings ($\overline{wage}_o$) and an assumed value for β, this provides a system of O equations that can be solved for unique values of the O unobserved human capital returns (ϕ_o).

Using these solutions for human capital returns (ϕ_o), the assumed values of β and θ, and the expressions for equilibrium human capital investments (A16) and the choice probabilities (A21), we can recover an adjusted measure of the relative effectiveness of workers in performing tasks for each sector and occupation ($\mathbb{A}_{so}$) from observed employment shares (λ_{so}) relative to their geometric mean:

$$\frac{\mathbb{A}_{so}}{\left[\prod_{s=1}^{S}\prod_{o=1}^{O}\mathbb{A}_{so}\right]^{\frac{1}{O}}} = \frac{\lambda_{so}/\left[\prod_{s=1}^{S}\prod_{o=1}^{O}\lambda_{so}\right]^{\frac{1}{O}}}{\mathbb{B}_{so}/\left[\prod_{s=1}^{S}\prod_{o=1}^{O}\mathbb{B}_{so}\right]^{\frac{1}{O}}}, \tag{A35}$$

where $\mathbb{B}_{so} = (1-\ell_o)^{\theta/\beta}\,\ell_o^{\theta\phi_o}$ captures the contribution of human capital investments to employment shares and $\mathbb{A}_{so} = T_{so}w_o^{\theta}\bar{h}_o^{\theta}$ captures the average effectiveness of workers in performing tasks within each sector and occupation (T_{so}), human capital productivity ($\bar{h}_o$), and the wage per effective unit of labor (w_o). We use the solutions for unobserved human capital returns (ϕ_o) from (A100) and adjusted task effectiveness ($\mathbb{A}_{so}$) from (A101) to quantify the relative importance of these two different mechanisms for explaining the observed changes in employment shares and average earnings in the data, as discussed in Section 5.3 of the paper and Section A6.3 of this web appendix.

A3 Multiple Worker Types Extension

In this section, we consider an extension of our baseline model that incorporates multiple worker types with different *ex ante* observable characteristics (e.g. age, gender and general schooling), as observed in the data. Worker types are indexed by $x \in \{1, \ldots, X\}$ and are imperfect substitutes in production. Each worker type x consists of a continuum of people ($\bar{L}^x$) who can choose to work in a number of occupations indexed by O^x. Human capital for each worker type in each occupation depends on raw worker ability and investments in human capital accumulation for that occupation. Human capital for each worker type in each occupation is used to produce final goods in a number of sectors indexed by $s = \{1, \ldots, S\}$.

A3.1 Preferences and Technology

A person i of type x with consumption C_i^x and leisure time $1-\ell_i^x$ obtains utility:

$$U_i^x = (C_i^x)^{\beta}(1-\ell_i^x), \qquad \beta > 0, \tag{A36}$$

where C_i^x is a consumption index; ℓ_i^x represents investments in human capital accumulation; and β parameterizes the tradeoff between consumption and the accumulation of human capital. The consumption index (C_i^x) is itself a Cobb-Douglas function of consumption of tradeable goods (C_{Mi}^x) and a non-tradeable good (C_{Ni}^x) that we interpret as housing:[6]

$$C_i^x = \left(\frac{C_{Mi}^x}{\alpha}\right)^{\alpha}\left(\frac{C_{Ni}^x}{1-\alpha}\right)^{1-\alpha}, \qquad 0 < \alpha < 1, \tag{A37}$$

[6]For empirical evidence in support of the constant housing expenditure share implied by this Cobb-Douglas functional form, see Davis and Ortalo-Magné (2011).

where housing is assumed to be in inelastic supply $\bar{N}$.

The tradeables consumption index (C^x_M) is a constant elasticity of substitution (CES) function of consumption of a number of sectors (C^x_{Ms}) indexed by $s \in \{0, \ldots, S\}$:

$$C^x_{Mi} = \left[\sum_{s=1}^{S} \left(\zeta_s C^x_{Mis} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \qquad \sum_{s=1}^{S} \zeta_s = 1, \tag{A38}$$

where ζ_s controls the strength of relative preferences for sector s; we assume for simplicity that preferences are the same across worker types; and σ is the elasticity of substitution between sectors. Output in each tradeable sector (Y_{Ms}) is a constant elasticity of substitution (CES) function of labor inputs of each worker type (H^x_s):

$$Y_{Ms} = \left[\sum_{x=1}^{X} \left(\vartheta^x_s H^x_s \right)^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}}, \qquad \sum_{x=1}^{X} \vartheta^x_s = 1 \; \forall \; s, \tag{A39}$$

where ϑ^x_s controls the relative productivity of each worker type in sector s; ρ is the elasticity of substitution between worker types x within sector s; and goods market clearing requires that output of each good equals the sum of consumption of that good across all individuals within each type and across types: $Y_{Ms} = \sum_x \sum_i C^x_{Mis}$.[7] The labor inputs of each worker type (H^x_s) depend on the human capital of workers of that type within each occupation (H^x_{so}):

$$H^x_s = \left[\sum_{o=1}^{O^x} \left(\xi^x_{so} H^x_{so} \right)^{\frac{\kappa-1}{\kappa}} \right]^{\frac{\kappa}{\kappa-1}}, \qquad \sum_{o=1}^{O^x} \xi^x_{so} = 1 \; \forall \; s, \tag{A40}$$

where ξ^x_{so} controls the relative productivity of worker type x in occupation o in sector s; this occupation o is not employed in sector s for worker type x if $\xi^x_{so} = 0$; κ is the elasticity of substitution between occupations o within worker type x and sector s; and we allow the number of occupations O^x to vary across worker types.[8]

Workers of each type choose an occupation and acquire human capital for that occupation, such that each occupation corresponds to a separate labor market for each worker type. Workers of a given type x within a given occupation o are mobile across sectors, which implies that the wage per effective unit of labor for workers of that type within that occupation (w^x_o) is the same across sectors. The choice of sector s for a worker i of type x within occupation o depends on idiosyncratic realizations for effective units of labor (ability z^x_{iso}). The choice of occupation o for a worker of type x depends on these realizations for idiosyncratic ability, the wage per effective unit of labor for each occupation for that worker type, and the rate of return to human capital investments for each occupation for that worker type.

Each person i of worker type x works one unit of time in her chosen occupation $o \in \{1, \ldots, O^x\}$. Another unit of time is divided between leisure ($1 - \ell^x_{io}$) and human capital accumulation (ℓ^x_{io}). The

[7] We assume for simplicity that labor is the sole factor of production, but the analysis can be extended to incorporate other factors of production, such as capital or land. We also assume for simplicity that ρ takes the same value across sectors, but it is straightforward to allow this elasticity to differ across sectors.

[8] We assume for simplicity that κ takes the same value across sectors, but it is also straightforward to allow this elasticity to differ across sectors.

production function for human capital for worker type x in occupation o is:

$$h^x_{io}\left(\ell^x_{io}\right) = \bar{h}^x_o \left(\ell^x_{io}\right)^{\phi^x_o}, \tag{A41}$$

where the parameter $\bar{h}^x_o > 0$ captures the productivity of human capital investments; the parameter $\phi^x_o > 0$ determines the rate of return to human capital accumulation; both parameters can differ across occupations o and worker types x.

Human capital for each worker type, sector and occupation (H^x_{so}) equals the fraction of workers of that type who choose that sector and occupation (λ^x_{so}) times average human capital of workers of that type conditional on choosing that sector and occupation times the measure of workers of that type in the economy ($\bar{L}^x$):

$$H^x_{so} = \lambda^x_{so}\, \mathbb{E}\left[h^x_o z^z \mid \text{Person of type } x \text{ chooses } s \text{ and } o\right] \bar{L}^x. \tag{A42}$$

where average human capital depends on both human capital accumulation (h^x_o) and ability (z^x).

The income of a worker of a given type depends on the wage per effective unit of labor for that worker type in her chosen occupation (w^x_o), her accumulated human capital (h^x_{io}) and her idiosyncratic ability (z^x_{iso}) for her chosen sector and occupation:

$$\Omega^x_i = w^x_o h^x_{io} z^x_{iso} = w^x_o \bar{h}^x_o \left(\ell^x_{io}\right)^{\phi^x_o} z^x_{iso}. \tag{A43}$$

The timing of decisions is as follows. First, each person i of worker type x observes her realizations of idiosyncratic ability (z^x_{iso}) and chooses a sector s and occupation o, taking wages (w^x_o) as given. Second, she chooses her optimal human capital investment in her chosen occupation (ℓ^x_{io}), given the trade-off between goods consumption and human capital investments in utility (A36) and the human capital accumulation technology (A41). Third, she makes her optimal choices for overall goods consumption (C^x_{Mi}), consumption of housing (C^x_{Ni}), and goods consumption for each sector (C^x_{Mis}), given observed prices (P_{Ms}) and her income (Ω^x_i), as determined by wages and human capital investments (A43).

An equilibrium in this economy is a set of allocations of consumption, production, human capital investments and choices of sector and occupation for each worker type $\{C^x_i, C^x_{Mi}, C^x_{Ni}, C^x_{Msi}, Y^x_{Msi}, \ell^x_{io}, \lambda^x_{so}\}$ and a set of prices $\{P_M, P_N, P_{Ms}, w^x_o\}$, such that individuals of each type choose consumption, human capital investments, sector and occupation to maximize utility; firms choose inputs of human capital to maximize profits; zero profits are made if a good is produced; and the markets for goods, labor and housing clear. We use the timing of decisions and the structure of the model to solve for equilibrium recursively. First, we characterize equilibrium consumption and production as a function of human capital investments and choice of sector and occupation for each worker type. Second, we determine optimal human capital investments as a function of choice of sector and occupation for each worker type. Third, we solve for the optimal choice of sector and occupation for each worker type.

A3.2 Consumption Decisions

Given human capital investments and a choice of sector and occupation for each worker type, the characterization of consumption decisions is straightforward. The Cobb-Douglas functional form (A37) implies that each person allocates constant shares of income to consumption of goods and housing: $C^x_{Mi} = \alpha\Omega^x_i/P_M$ and $C^x_{Ni} = (1-\alpha)\Omega^x_i/P_N$. Using these results in (A36), the utility function can be written in terms of income, the prices of goods consumption and housing, and investment in human capital accumulation:

$$U^x_i = \left(\frac{w^x_o \bar{h}^x_o \left(\ell^x_{io}\right)^{\phi^x_o} z^x_{iso}}{P^\alpha_M P^{1-\alpha}_N}\right)^\beta \left(1-\ell^x_{io}\right). \tag{A44}$$

Using the CES functional form of goods consumption (A38), the share of goods expenditure allocated to each sector (μ_{Ms}) depends on prices (P_{Ms}) and preference weights (ζ_s):

$$\mu_{Ms} = \frac{P_{Ms}C_{Ms}}{\sum_{k=1}^S P_{Mk}C_{Mk}} = \frac{\left(P_{Ms}/\zeta_s\right)^{1-\sigma}}{\sum_{k=1}^S \left(P_{Mk}/\zeta_k\right)^{1-\sigma}}, \tag{A45}$$

and the overall price index (P_M) dual to the goods consumption index (A38) can be written in terms of prices (P_{Ms}) and preference weights (ζ_s) for each sector:

$$P_M = \left[\sum_{s=1}^S \left(\frac{P_{Ms}}{\zeta_s}\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}. \tag{A46}$$

Using the CES production technology (A39), the share of each worker type in the wage bill in each sector depends on the unit cost function for each worker type (ω^x_s) and productivities (ϑ^x_s):

$$\frac{\omega^x_s H^x_s}{\sum_{a=1}^X \omega^a_s H^a_s} = \frac{\left(\omega^x_s/\vartheta^x_s\right)^{1-\rho}}{\sum_{a=1}^X \left(\omega^a_s/\vartheta^a_s\right)^{1-\rho}}. \tag{A47}$$

Perfect competition implies that prices in each sector (P_{Ms}) equal unit costs, which can be expressed as the following function of unit cost function for each worker type (ω^x_s) and productivities (ϑ^x_s):

$$P_{Ms} = \left[\sum_{x=1}^X \left(\frac{\omega^x_s}{\vartheta^x_s}\right)^{1-\rho}\right]^{\frac{1}{1-\rho}}. \tag{A48}$$

Using the CES technology (A40) for labor input of each type, the share of each occupation in the wage bill for each worker type in each sector depends on occupational wages for that worker type (w^x_o) and productivities (ξ^x_{so}):

$$\frac{w^x_o H^x_{so}}{\sum_{m=1}^{O^x} w^x_m H^x_{so}} = \frac{\left(w^x_o/\xi^x_{so}\right)^{1-\kappa}}{\sum_{m=1}^{O^x} \left(w^x_m/\xi^x_{sm}\right)^{1-\kappa}}. \tag{A49}$$

Cost minimization implies that the unit cost for each worker type in each sector (ω^x_s) can be expressed as the following function of occupational wages for that worker type (w^x_o) and productivities (ξ^x_{so}):

$$\omega^x_s = \left[\sum_{o=1}^{O^x} \left(\frac{w^x_o}{\xi^x_{so}}\right)^{1-\kappa}\right]^{\frac{1}{1-\kappa}}. \tag{A50}$$

The wage (w^x_o) for occupation o and worker type x is determined by the requirement that the total income of all workers of that type in that occupation is equal to the total payments to workers of that type in that occupation:

$$w^x_o \sum_{s=1}^S H^x_{so} = \sum_{s=1}^S \frac{\left(\omega^x_s/\vartheta^x_s\right)^{1-\rho}}{\sum_{a=1}^X \left(\omega^a_s/\vartheta^a_s\right)^{1-\rho}} \frac{\left(w^x_o/\xi^x_{so}\right)^{1-\kappa}}{\sum_{m=1}^{O^x} \left(w^x_m/\xi^x_{sm}\right)^{1-\kappa}} \mu_{Ms} E. \tag{A51}$$

where the right-hand side equals the sum across sectors of the share of worker type x in each sector's wage bill times the share of occupation o in the wage bill for worker type x in each sector times total labor payments for each sector (which equal the sector's expenditure share (μ_{Ms}) times total expenditure (E)). Finally, the price of housing is determined by housing market clearing:

$$P_N = \frac{(1-\alpha)\Omega}{\bar{N}}, \tag{A52}$$

where recall that $\bar{N}$ is the inelastic supply of housing and aggregate income (Ω) is the sum of the income of all persons i of all types x (Ω_i^x). The larger the measure of people in the economy, and the higher their income relative to the supply of housing, the higher the price of housing.

A3.3 Human Capital Investments

Given a choice of sector s and occupation o, wages (w_o^x), realizations of idiosyncratic ability (z_{iso}^x), and prices of tradeable (P_M) and non-tradeable (P_N) goods, each individual i of each worker type x chooses her human capital investment (ℓ_{io}^x) to maximize her utility (A44):

$$\max_{\ell_{io}^x} \left\{ \left(\frac{w_o^x \bar{h}_o^x \left(\ell_{io}^x\right)^{\phi_o^x} z_{iso}^x}{P_M^\alpha P_N^{1-\alpha}} \right)^\beta \left(1-\ell_{io}^x\right) \right\}. \tag{A53}$$

The first-order condition to this problem yields the equilibrium human capital investment:

$$\ell_o^{x*} = \frac{1}{1+\frac{1}{\beta\phi_o^x}}, \tag{A54}$$

which only varies across occupations o and worker types x and not across individuals i within each occupation and worker type. Hence, from now onwards, we suppress the individual subscript i, unless otherwise indicated.

Equilibrium human capital investments depend solely on β (the tradeoff between consumption and the accumulation of human capital) and ϕ_o^x (the productivity of human capital accumulation in the worker's chosen occupation). Other forces do not affect human capital investments, because they have the same effect on the return and opportunity cost of time spent accumulating human capital.

A3.4 Sector and Occupation Choice

Given wages (w_o^x), realizations of idiosyncratic ability (z_{iso}^x), and prices of tradeable (P_M) and non-tradeable (P_N) goods, each individual of each worker type chooses her sector and occupation to maximize her utility. We model individual ability following McFadden (1974) and Eaton and Kortum (2002). Each individual i of worker type x draws ability for each sector s and occupation o (z_{iso}^x) from an independent Fréchet distribution:

$$F_{so}^x(z) = e^{-T_{so}^x z^{-\theta^x}}, \qquad \theta^x > 1. \tag{A55}$$

The Fréchet scale parameter T_{so}^x determines the average ability of workers of type x in performing tasks in sector s and occupation o. The Fréchet shape parameter θ^x determines the dispersion of the ability of workers of type x in performing tasks across sectors and occupations. A reduction in θ^x corresponds to increased dispersion of idiosyncratic ability and greater scope for workers of a given type to specialize according comparative advantage across sectors and occupations.[9]

From utility (A44), the following transformation of utility is linear in worker ability:

$$v^x = (U^x)^{1/\beta} = \bar{w}_o^x z, \qquad \bar{w}_o^x = \frac{w_o^x \left(1-\ell_o^x\right)^{1/\beta} \left(\ell_o^x\right)^{\phi_o^x} \bar{h}_o^x}{P_M^{\alpha} P_N^{1-\alpha}}. \tag{A56}$$

Using this monotonic relationship between utility and worker ability, the distribution of utility across workers of a given type x within each sector s and occupation o inherits a Fréchet distribution:

$$F_{so}^x(v) = e^{-\Phi_{so}^x v^{-\theta^x}}, \qquad \Phi_{so}^x = T_{so}^x \left(\bar{w}_o^x\right)^{\theta^x}. \tag{A57}$$

Each worker of each type chooses their occupation and sector to maximize their utility. Note that the maximum of Fréchet distributed random variables also has a Fréchet distribution. Therefore the distribution of utility across all sectors and occupations for each worker type is given by:

$$F^x(v) = e^{-\Phi^x v^{-\theta^x}}, \qquad \Phi^x = \sum_{s=1}^{S} \sum_{o=1}^{O} T_{so}^x \left(\bar{w}_o^x\right)^{\theta^x}, \tag{A58}$$

where the Fréchet functional form implies that the distribution of utility conditional on choosing a sector and occupation is the same for each sector and occupation pair for a given worker type and is equal to the distribution of utility across all sectors and occupations (A58) for that worker type.

Sector and occupation choice for each worker type takes a similar form as in our baseline specification above. The model again yields two sufficient statistics for employment shares and average wages, which now differ across worker types.

Proposition 3 ***(Multiple Worker Types)*** *Let λ_{so}^x denote the fraction of people of type x who choose to work in sector s and occupation o. Let λ_o^x denote the fraction of people of type x who choose to work in occupation o. Let $\overline{wage}_o^x$ denote average earnings for people of type x who choose to work in occupation o. The sufficient statistics for employment shares and average wages are:*

$$\lambda_{so}^x = \frac{\Psi_{so}^x}{\Psi^x}, \qquad \lambda_o^x = \frac{\Psi_o^x}{\Psi^x}, \tag{A59}$$

$$\Psi_o^x = \sum_{s=1}^{S} \Psi_{so}^x, \qquad \Psi^x = \sum_{o=1}^{O} \sum_{s=1}^{S} \Psi_{so}^x, \qquad \Psi_{so}^x = T_{so}^x \left(w_o^x\right)^{\theta} \left(1-\ell_o^x\right)^{\theta^x/\beta} \left(\ell_o^x\right)^{\theta^x \phi_o^x} \left(\bar{h}_o^x\right)^{\theta^x}.$$

$$\overline{wage}_o^x = \mathbb{E}\left[w_o^x h_o^x z_o^x\right] = \gamma^x \left(1-\ell_o^x\right)^{-1/\beta} \left(P_M^{\alpha} P_N^{1-\alpha}\right) \left(\Phi^x\right)^{1/\theta^x}, \tag{A60}$$

where $\gamma^x = \Gamma\left(\frac{\theta^x - 1}{\theta^x}\right)$ and $\Gamma(\cdot)$ is the Gamma function.

[9] Although we assume that ability is drawn independently for each sector and occupation for workers of a given type, the parameter T_{so}^x induces a correlation in ability among workers of a given type within the same sector and occupation. While we focus on the independent Fréchet distribution for simplicity, it is straightforward to consider the multivariate Fréchet distribution, which allows for correlation in the ability draws of individual workers of a given type across sectors and occupations.

Proof. We begin with the derivation of the employment shares $\{\lambda_{so}^x, \lambda_o^x\}$. The probability that a worker of type x chooses sector s and occupation o is:

$$
\begin{aligned}
\lambda_{so}^x =& \Pr\left[v_{so}^x \geq \max\{v_{km}^x\}; \forall k, m\right], \\
=& \int_0^\infty \prod_{k\neq s} F_{ko}^x(v) \left[\prod_{k=1}^{S}\prod_{m\neq o} F_{km}^x(v)\right] f_{so}^x(v)dv, \\
=& \int_0^\infty \prod_{k\neq s} e^{-\Phi_{ko}^x v^{-\theta^x}} \left[\prod_{k=1}^{S}\prod_{m\neq o} e^{-\Phi_{km}^x v^{-\theta^x}}\right] \theta^x \Phi_{so}^x v^{-(\theta^x+1)} e^{-\Phi_{so}^x v^{-\theta^x}} dv, \\
=& \int_0^\infty \prod_{k=1}^{S}\prod_{m=1}^{O} \theta^x \Phi_{so}^x v^{-(\theta^x+1)} e^{-\Phi_{km}^x v^{-\theta^x}} dv, \\
=& \int_0^\infty \theta^x \Phi_{so}^x v^{-(\theta^x+1)} e^{-\Phi^x} v^{-\theta^x} dv.
\end{aligned}
$$

Note that:

$$
\frac{d}{dv}\left[-\frac{1}{\Phi^x} e^{-\Phi^x v^{-\theta^x}}\right] = \theta^x v^{-(\theta^x+1)} e^{-\Phi^x}.
$$

Therefore:

$$
\lambda_{so}^x = \left[-\frac{\Phi_{so}^x}{\Phi^x} e^{-\Phi^x v^{-\theta^x}}\right]_0^\infty,
$$

which becomes:

$$
\lambda_{so}^x = \frac{\Phi_{so}^x}{\Phi^x} = \frac{\Psi_{so}^x}{\Psi^x},
$$

where

$$
\frac{\Phi_{so}^x}{\Phi^x} = \frac{\frac{T_{so}^x (w_o^x)^{\theta^x} (1-\ell_o^x)^{\theta^x/\beta} (\ell_o^x)^{\theta^x \phi_o^x} \left(\bar{h}_o^x\right)^{\theta^x}}{P_M^{\alpha\theta^x} P_N^{(1-\alpha)\theta^x}}}{\sum_{k=1}^{S}\sum_{m=1}^{O^x} \frac{T_{km}^x (w_m^x)^{\theta^x} (1-\ell_m^x)^{\theta^x/\beta} (\ell_m^x)^{\theta^x \phi_m^x} \left(\bar{h}_m^x\right)^{\theta^x}}{P_M^{\alpha\theta^x} P_N^{(1-\alpha)\theta^x}}},
$$

which simplifies to:

$$
\frac{\Psi_{so}^x}{\Psi^x} = \frac{T_{so}^x \left(w_o^x\right)^{\theta^x} \left(1-\ell_o^x\right)^{\theta^x/\beta} \left(\ell_o^x\right)^{\theta^x\phi_o^x} \left(\bar{h}_o^x\right)^{\theta^x}}{\sum_{k=1}^{S}\sum_{m=1}^{O^x} T_{km}^x \left(w_m^x\right)^{\theta^x} \left(1-\ell_m^x\right)^{\theta^x/\beta} \left(\ell_m^x\right)^{\theta^x\phi_m^x} \left(\bar{h}_m^x\right)^{\theta^x}}.
$$

Summing across sectors s, we obtain the probability that a worker of type x chooses occupation o:

$$
\lambda_o^x = \sum_{s=1}^{S} \lambda_{so}^x = \frac{\Psi_o^x}{\Psi^x}, \qquad \Psi_o^x = \sum_{s=1}^{S} \Psi_{so}^x = \sum_{s=1}^{S} T_{so}^x \left(w_o^x\right)^{\theta^x} \left(1-\ell_o^x\right)^{\theta^x/\beta} \left(\ell_o^x\right)^{\theta^x\phi_o^x} \left(\bar{h}_o^x\right)^{\theta^x}.
$$

We next derive average earnings ($\overline{wage}_o^x$). Average earnings in the chosen occupation of a worker of a given type depend on the wage per effective unit of labor, her human capital accumulation and her average ability conditional on choosing that occupation:

$$
\overline{wage}_o^x = \mathbb{E}\left[w_o^x h_o^x z \mid \text{type } x \text{ chooses } o\right]. \tag{A61}
$$

To derive the distribution of ability for a worker of a given type conditional on choosing an occupation, note that an implication of the Fréchet distribution is that the distribution of utility for a worker of a given type conditional on choosing a sector and occupation is the same for all sectors and occupations, and equal to the distribution of utility for that worker type across all sectors and occupations (A58), reproduced below:

$$F^x(v) = e^{-\Phi^x v^{-\theta^x}}, \qquad \Phi^x = \sum_{k=1}^{S} \sum_{m=1}^{O^x} T^x_{km} (\bar{w}^x_m)^{\theta^x}.$$

Using the monotonic relationship between utility and ability for workers of a given type:

$$v^x = \bar{w}^x_o z,$$

we obtain the distribution of ability conditional on choosing an occupation for workers of a given type:

$$F^x(z) = e^{-\Phi^{x*} z^{-\theta^x}}, \qquad \Phi^{x*} = \sum_{k=1}^{S} \sum_{m=1}^{O^x} T^x_{km} (\bar{w}^x_m / \bar{w}^x_o)^{\theta^x}.$$

Therefore expected ability in a worker's chosen occupation for a given worker type is:

$$\mathbb{E}[z \mid \text{type } x \text{ chooses } o] = \int_0^\infty \theta^x \Phi^{x*} z^{-\theta^x} e^{-\Phi^{x*} z^{\theta^x}} dz.$$

Defining the following change of variables:

$$y^x = \Phi^{x*} z^{-\theta^x}, \qquad dy^x = \theta^x \Phi^{x*} z^{-(\theta^x+1)} dz,$$

expected ability for a given worker type can be re-written as:

$$\mathbb{E}[z \mid \text{type } x \text{ chooses } o] = \int_0^\infty (y^x)^{-1/\theta^x} (\Phi^{x*})^{1/\theta^x} e^{-y^x} dy^x,$$

which yields:

$$\mathbb{E}[z \mid \text{type } x \text{ chooses } o] = \gamma^x (\Phi^{x*})^{1/\theta^x}, \qquad \gamma^x = \Gamma\left(\frac{\theta^x - 1}{\theta^x}\right), \tag{A62}$$

where $\Gamma(\cdot)$ is the gamma function. Now note that:

$$\lambda^x_o = \frac{\sum_k T^x_{ko} (\bar{w}^x_o)^{\theta^x}}{\sum_{k=1}^{S} \sum_{m=1}^{O^x} T^x_{km} (\bar{w}^x_m)^{\theta^x}}, \tag{A63}$$

which can be written as:

$$\lambda^x_o = \frac{\sum_k T^x_{ko}}{\sum_{k=1}^{S} \sum_{m=1}^{O^x} T^x_{km} (\bar{w}^x_m / w^x_o)^{\theta^x}} = \frac{\sum_k T^x_{ko}}{\Phi^{x*}}.$$

Therefore expected ability in a worker's chosen occupation for a given worker type (A62) can be re-expressed as:

$$\mathbb{E}[z \mid \text{type } x \text{ chooses } o] = \gamma^x \left(\frac{\sum_{k=1}^{S} T^x_{ko}}{\lambda^x_o}\right)^{1/\theta^x}. \tag{A64}$$

It follows that average human capital in a worker's chosen occupation for a given worker type is:

$$\mathbb{E}[h^x_o z \mid \text{type } x \text{ chooses } o] = \gamma^x \bar{h}^x_o (\ell^x_o)^{\phi^x_o} \left(\frac{\sum_{k=1}^{S} T^x_{ko}}{\lambda^x_o}\right)^{1/\theta^x}. \tag{A65}$$

Using the occupational choice probability (A63), average human capital in a worker's chosen occupation for a given worker type can be re-written as:

$$\mathbb{E}\left[h_o z \mid \text{type } x \text{ chooses } o\right] = \gamma^x \bar{h}_o^x \left(\ell_o^x\right)^{\phi_o^x} \frac{1}{\bar{w}_o^x} \left(\sum_{k=1}^{S} \sum_{m=1}^{O^x} T_{km}^x \left(\bar{w}_m^x\right)^{\theta^x}\right)^{1/\theta^x}. \tag{A66}$$

Using the definition of $\bar{w}_o^x$ in (A56), it follows that average income in a worker's chosen occupation for a given worker type is:

$$\mathbb{E}\left[w_o^x h_o^x z \mid \text{type } x \text{ chooses } o\right] = \gamma^x \left(1 - \ell_o^x\right)^{-1/\beta} P_M^\alpha P_N^{1-\alpha} \left(\sum_{k=1}^{S} \sum_{m=1}^{O^x} T_{km}^x \left(\bar{w}_m^x\right)^{\theta^x}\right)^{1/\theta^x}. \tag{A67}$$

Therefore relative average income for a given worker type only varies across occupations because of variation in human capital investments ℓ_o^x:

$$\frac{\mathbb{E}\left[w_o^x h_o^x z \mid x \text{ chooses } o\right]}{\mathbb{E}\left[w_m^x h_m^x z \mid x \text{ chooses } m\right]} = \left(\frac{1 - \ell_o^x}{1 - \ell_m^x}\right)^{-1/\beta}. \tag{A68}$$

■

Therefore this extension to multiple worker types preserves the key properties of the baseline model. Differences across worker types in the average effectiveness of performing tasks within each sector and occupation (T_{so}^x) generate variation across worker types in employment shares (λ_{so}^x). In contrast, the return to human capital investments in each occupation (ϕ_o^x) affects both employment shares (λ_{so}^x) and average occupational earnings ($\overline{wage}_o^x$). To the extent that new technologies complement or substitute for workers of a given type in performing tasks in each sector and occupation (raising or reducing T_{so}^x), they increase or decrease employment shares respectively. To the extent that these new technologies also complement or substitute for human capital investments (ϕ_o^x), they raise or reduce both employment shares and average wages respectively.

A4 Multi-region Extension

In this section, we consider an extension of our baseline model that incorporates multiple regions (e.g. urban versus rural areas), as observed in the data. Although for simplicity we return to the case of a single worker type, the two extensions can be combined to allow for both multiple worker types and multiple regions. The economy consists of a continuum of people ($\bar{L}$) who can choose a region $r \in \{1, \ldots, R\}$, an occupation $o \in \{1, \ldots, O\}$, and a sector $s \in \{1, \ldots, S\}$. Human capital for each occupation depends on raw worker ability and investments in human capital accumulation for that occupation. People choose an occupation based on the wage and cost of investing in human capital for that occupation. Human capital for each occupation is used to produce final goods in each sector in each region. People choose a region based on the wage and the cost of living in that region. We allow some occupations not be employed in some sectors and some sectors not to be produced in some regions.

A4.1 Preferences and Technology

A person i in region r with consumption C_{ir} and leisure time $1-\ell_{ir}$ obtains utility:

$$U_{ir} = C_{ir}^{\beta}\left(1-\ell_{ir}\right), \qquad \beta > 0, \tag{A69}$$

where C_{ir} is a consumption index; ℓ_{ir} represents investments in human capital accumulation; and β parameterizes the tradeoff between consumption and the accumulation of human capital. The consumption index (C_{ir}) is itself a Cobb-Douglas function of consumption of tradeable goods (C_{Mir}) and a non-tradeable good (C_{Nir}) that we interpret as housing:[10]

$$C_{ir} = \left(\frac{C_{Mir}}{\alpha}\right)^{\alpha}\left(\frac{C_{Nir}}{1-\alpha}\right)^{1-\alpha}, \qquad 0<\alpha<1, \tag{A70}$$

where housing is assumed to be in inelastic supply $\bar{N}_r$ and provides a force for the dispersion of economic activity across regions.

The tradeables consumption index (C_{Mir}) is a constant elasticity of substitution (CES) function of consumption of a number of sectors (C_{Mirs}) indexed by $s \in \{0,\ldots,S\}$:

$$C_{Mir} = \left[\sum_{s=1}^{S}\left(\zeta_s C_{Mirs}\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}, \qquad \sum_{s=1}^{S}\zeta_s = 1, \tag{A71}$$

where ζ_s controls the strength of relative preferences for sector s; we assume for simplicity that preferences are the same across regions; and σ is the elasticity of substitution between sectors. Output in each region and sector (Y_{Mrs}) is a constant elasticity of substitution (CES) function of the human capital of workers from each occupation within that region and sector (H_{rso}):

$$Y_{Mrs} = \left[\sum_{o=1}^{O}\left(\xi_{rso} H_{rso}\right)^{\frac{\kappa-1}{\kappa}}\right]^{\frac{\kappa}{\kappa-1}}, \qquad \sum_{o=1}^{O}\xi_{rso} = 1 \;\forall\, s, \tag{A72}$$

where ξ_{rso} controls the relative productivity of occupation o in region r and sector s; κ is the elasticity of substitution between occupations o within sector s; and goods market clearing requires that output of each good equals the sum of consumption of that good across all individuals and regions: $Y_{Ms} = \sum_r \sum_i C_{Mirs}$.[11]

Workers choose an occupation for which to acquire human capital and a region in which to live, such that each occupation and region corresponds to a separate labor market. Workers within a given occupation o and region r are mobile across sectors, which implies that the wage per effective unit of labor for workers within that occupation and region (w_{ro}) is the same across sectors. The choice of sector s for a worker i within occupation o and region r depends on idiosyncratic realizations for effective units of

[10] For empirical evidence in support of the constant housing expenditure share implied by this Cobb-Douglas functional form, see Davis and Ortalo-Magné (2011).

[11] We assume for simplicity that labor is the sole factor of production, but the analysis can be extended to incorporate other factors of production, such as capital or land. We also assume for simplicity that κ takes the same value across regions, but it is straightforward to allow this elasticity to differ across regions.

labor (ability z_{irso}). The choice of occupation o and region r depends on these realizations for idiosyncratic ability, the wage per effective unit of labor for each occupation and region, and the rate of return to human capital investments for each occupation.

Each person i works one unit of time in her chosen occupation $o \in \{1, \ldots, O\}$. Another unit of time is divided between leisure $(1 - \ell_{iro})$ and human capital accumulation (ℓ_{iro}). The production function for human capital in occupation o is:

$$h_{iro}\left(\ell_{iro}\right) = \bar{h}_o \ell_{iro}^{\phi_o}, \tag{A73}$$

where the parameter $\bar{h}_o > 0$ captures the productivity of human capital investments; the parameter $\phi_o > 0$ determines the rate of return to human capital accumulation; and both parameters can differ across occupations o.

Human capital for each region, sector and occupation (H_{rso}) equals the fraction of agents who choose that region, sector and occupation (λ_{rso}) times average human capital conditional on choosing that region, sector and occupation times the measure of agents in the economy $(\bar{L})$:

$$H_{rso} = \lambda_{rso} \, \mathbb{E}\left[h_{ro} z \mid \text{Person chooses } r, \ s \text{ and } o\right] \bar{L}. \tag{A74}$$

where average human capital depends on both human capital accumulation (h_{ro}) and ability (z).

Each person's income depends on the wage per effective unit of labor for her chosen region and occupation (w_{ro}), her accumulated human capital (h_{iro}) and her idiosyncratic ability (z_{irso}) for her chosen region, sector and occupation:

$$\Omega_{ir} = w_{ro} h_{iro} z_{irso} = w_{ro} \bar{h}_o \ell_{iro}^{\phi_o} z_{irso}. \tag{A75}$$

The timing of decisions is as follows. First, each person i observes her realizations of idiosyncratic ability (z_{irso}) and chooses a region r, sector s and occupation o, taking wages (w_{ro}) as given. Second, she chooses her optimal human capital investment in her chosen occupation (ℓ_{iro}), given the trade-off between goods consumption and human capital investments in utility (A1) and the human capital accumulation technology (A5). Third, she makes her optimal choices for overall goods consumption (C_{Mir}), consumption of housing (C_{Nir}), and goods consumption for each sector (C_{Mirs}), given observed prices (P_{Mrs}) and her income (Ω_{ir}), as determined by wages and human capital investments (A75).

An equilibrium in this economy is a set of allocations of consumption, production, human capital investments and choices of region, sector and occupation $\{C_{ir}, C_{Mir}, C_{Nir}, C_{Mirs}, Y_{Mrs}, \ell_{iro}, \lambda_{rso}\}$ and a set of prices $\{P_{Mr}, P_{Nr}, P_{Mrs}, w_{ro}\}$, such that individuals choose consumption, human capital investments as well as region, sector and occupation to maximize utility; firms choose inputs of human capital to maximize profits; zero profits are made if a good is produced; and the markets for goods, labor and housing clear. We use the timing of decisions and the structure of the model to solve for equilibrium recursively. First, we characterize equilibrium consumption and production as a function of human capital investments and choice of region, sector and occupation. Second, we determine optimal human capital investments as a function of choice of region, sector and occupation. Third, we solve for the optimal choice of region, sector and occupation.

A4.2 Consumption Decisions

Given an individual's human capital investments and her choice of region, sector and occupation, the characterization of consumption decisions is straightforward. The Cobb-Douglas functional form (A70) implies that each person allocates constant shares of income to consumption of goods and housing: $C_{Mir} = \alpha\Omega_{ir}/P_{Mr}$ and $C_{Nir} = (1-\alpha)\Omega_{ir}/P_{Nr}$. Using these results in (A69), the utility function can be written in terms of income, the prices of goods consumption and housing, and investment in human capital accumulation:

$$U_{ir} = \left(\frac{w_{ro}\bar{h}_o \ell_{iro}^{\phi_o} z_{irso}}{P_{Mr}^{\alpha} P_{Nr}^{1-\alpha}} \right)^{\beta} (1 - \ell_{iro}) \,. \tag{A76}$$

Using the CES functional form of goods consumption (A71), the share of goods expenditure allocated to each sector in each region (μ_{Mrs}) depends on prices (P_{Mrs}) and preference weights (ζ_s):

$$\mu_{Mrs} = \frac{P_{Mrs}C_{Mrs}}{\sum_{k=1}^{S} P_{Mrk}C_{Mrk}} = \frac{(P_{Mrs}/\zeta_s)^{1-\sigma}}{\sum_{k=1}^{S} (P_{Mrk}/\zeta_k)^{1-\sigma}}, \tag{A77}$$

and the overall price index (P_{Mr}) dual to the goods consumption index (A71) for each region can be written in terms of prices (P_{Mrs}) for each sector and region and preference weights (ζ_s) for each sector:

$$P_{Mr} = \left[\sum_{s=1}^{S} \left(\frac{P_{Mrs}}{\zeta_s} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}}. \tag{A78}$$

Using the CES production technology (A72), the share of each occupation in the wage bill in each sector and region depends on occupational wages in each region (w_{ro}) and productivities in each sector and region (ξ_{rso}):

$$\frac{w_{ro}H_{rso}}{\sum_{m=1}^{O} w_{rm}H_{rsm}} = \frac{(w_{ro}/\xi_{rso})^{1-\kappa}}{\sum_{m=1}^{O} (w_{rm}/\xi_{rsm})^{1-\kappa}}. \tag{A79}$$

Perfect competition implies that prices in each sector and region (P_{Mrs}) equal unit costs, which can be expressed as the following function of occupational wages in each region (w_{ro}) and productivities (ξ_{rso}) in each sector and region:

$$P_{Mrs} = \left[\sum_{o=1}^{O} \left(\frac{w_{ro}}{\xi_{rso}} \right)^{1-\kappa} \right]^{\frac{1}{1-\kappa}}. \tag{A80}$$

The occupational wage in each region (w_{ro}) is determined by the requirement that the total income of all workers within the occupation and region is equal to total payments to workers employed in that occupation and region:

$$w_{ro} \sum_{s=1}^{S} H_{rso} = \sum_{s=1}^{S} \frac{(w_{ro}/\xi_{rso})^{1-\kappa}}{\sum_{m=1}^{O} (w_{rm}/\xi_{rsm})^{1-\kappa}} P_{Mrs}Y_{Mrs}. \tag{A81}$$

Finally, the price of housing in each region is determined by housing market clearing:

$$P_{Nr} = \frac{(1-\alpha)\Omega_r}{\bar{N}_r}, \tag{A82}$$

where recall that $\bar{N}_r$ is the inelastic supply of housing and regional income (Ω_r) is the sum of income for each person i within that region r (Ω_{ir}). The larger the measure of people who locate in a region, and the higher their income relative to the supply of housing, the higher the price of housing in that region.

A4.3 Human Capital Investments

Given a choice of region r, sector s and occupation o, wages (w_{ro}), realizations of idiosyncratic ability (z_{irso}), and prices of tradeable (P_{Mr}) and non-tradeable (P_{Nr}) goods, each individual i chooses her human capital investment (ℓ_{iro}) to maximize her utility (A76):

$$\max_{\ell_{iro}} \left\{ \left(\frac{w_{ro}\bar{h}_o \ell_{iro}^{\phi_o} z_{irso}}{P_{Mr}^{\alpha} P_{Nr}^{1-\alpha}} \right)^{\beta} (1-\ell_{iro}) \right\}. \tag{A83}$$

The first-order conditions to this problem yield the equilibrium human capital investment:

$$\ell_o^* = \frac{1}{1+\frac{1}{\beta\phi_o}}, \tag{A84}$$

which only varies across occupations o and not across regions, sectors or individuals i within an occupation. From now onwards, we suppress the individual subscript i, unless otherwise indicated.

A4.4 Region, Sector and Occupation Choice

Given wages (w_{ro}), realizations of idiosyncratic ability (z_{irso}), and prices of tradeable (P_{Mr}) and non-tradeable (P_{Nr}) goods, each individual chooses her region, sector and occupation to maximize her utility. We model individual ability following McFadden (1974) and Eaton and Kortum (2002). Each individual i draws ability for each region r, sector s and occupation o (z_{irso}) from an independent Fréchet distribution:

$$F_{rso}(z) = e^{-T_{rso} z^{-\theta}}, \qquad \theta > 1. \tag{A85}$$

The Fréchet scale parameter T_{rso} determines the average ability of workers in performing tasks in region r, sector s and occupation o. The Fréchet shape parameter θ determines the dispersion of the ability of workers in performing tasks across regions, sectors and occupations. A reduction in θ corresponds to increased dispersion of idiosyncratic ability and greater scope for worker specialization according comparative advantage across regions, sectors and occupations.[12] The Fréchet scale parameter T_{rso} that determines average worker productivity can be either exogenous or endogenous to the concentration of

[12] Although we assume that ability is drawn independently for each region, sector and occupation, the parameter T_{rso} induces a correlation in ability among workers within the same region, sector and occupation. While we focus on the independent Fréchet distribution for simplicity, it is straightforward to consider the multivariate Fréchet distribution, which allows for correlation in the ability draws of individual workers across regions, sectors and occupations.

economic activity in a region (through agglomeration forces in the form of external economies of scale). In general, these external economies of scale can depend on the full vector of employment ($\boldsymbol{H}$) across regions, sectors and occupations ($T_{rso} = \Im_{rso}(\boldsymbol{H})$).

From utility (A76), the following transformation of utility is linear in worker ability:

$$v = U^{1/\beta} = \bar{w}_{ro} z, \qquad \bar{w}_{ro} = \frac{w_{ro} \left(1 - \ell_o\right)^{1/\beta} \ell_o^{\phi_o} \bar{h}_o}{P_{Mr}^{\alpha} P_{Nr}^{1-\alpha}}. \tag{A86}$$

Using this monotonic relationship between utility and worker ability, the distribution of utility across workers within each region r, sector s and occupation o inherits a Fréchet distribution:

$$F_{rso}\left(v\right) = e^{-\Phi_{rso} v^{-\theta}}, \qquad \Phi_{rso} = T_{rso} \bar{w}_{ro}^{\theta}. \tag{A87}$$

Each worker chooses their region, sector and occupation to maximize their utility. Note that the maximum of Fréchet distributed random variables also has a Fréchet distribution. Therefore the distribution of utility across all regions, sectors and occupations is given by:

$$F\left(v\right) = e^{-\Phi v^{-\theta}}, \qquad \Phi = \sum_{r=1}^{R} \sum_{s=1}^{S} \sum_{o=1}^{O} T_{rso} \bar{w}_{ro}^{\theta}, \tag{A88}$$

where the Fréchet functional form implies that the distribution of utility conditional on choosing a region, sector and occupation is the same for all combinations of region, sector and occupation and is equal to the distribution of utility across all regions, sectors and occupations (A88).

Each worker's choice of region, sector and occupation takes a similar form as her choice of sector and occupation in our baseline specification above. The model again yields two sufficient statistics for employment shares and average wages:

Proposition 4 ***(Multiple Regions)*** *Let λ_{rso} denote the fraction of people who choose to work in sector s, occupation o and region r. Let λ_{ro} denote the fraction of people who choose to work in occupation o in region r. Let $\overline{wage}_{ro}$ denote average earnings for people who choose to work in occupation o in region r. The sufficient statistics for employment shares and average wages are:*

$$\lambda_{rso} = \frac{\Phi_{rso}}{\Phi}, \qquad \lambda_{ro} = \frac{\Phi_{ro}}{\Phi}, \tag{A89}$$

$$\Phi_{ro} = \sum_{s=1}^{S} \Phi_{rso}, \qquad \Phi = \sum_{r=1}^{R} \sum_{o=1}^{O} \sum_{s=1}^{S} \Phi_{rso}, \qquad \Phi_{rso} = \frac{T_{rso} w_{ro}^{\theta} \left(1 - \ell_o\right)^{\theta/\beta} \ell_o^{\theta\phi_o} \bar{h}_o^{\theta}}{P_{Mr}^{\alpha} P_{Nr}^{1-\alpha}},$$

$$\overline{wage}_{ro} = \mathbb{E}\left[w_{ro} h_o z_{ro}\right] = \gamma \left(1 - \ell_o\right)^{-1/\beta} \left(P_{Mr}^{\alpha} P_{Nr}^{1-\alpha}\right) \Phi_r^{1/\theta}. \tag{A90}$$

where recall $\gamma = \Gamma\left(\frac{\theta-1}{\theta}\right)$ and $\Gamma\left(\cdot\right)$ is the Gamma function.

Proof. We begin with the derivation of the employment shares $\{\lambda_{rso}, \lambda_{ro}\}$. The probability that a worker chooses sector s and occupation o is:

$$\begin{aligned}
\lambda_{rso} =& \Pr\left[v_{rso} \geq \max\{v_{jkm}\}; \forall j,k,m\right], \\
=& \int_0^\infty \prod_{j\neq r}\prod_{k\neq s} F_{jko}(v)\left[\prod_{j=1}^{R}\prod_{k=1}^{S}\prod_{m\neq o} F_{jkm}(v)\right] f_{rso}(v)dv, \\
=& \int_0^\infty \prod_{j\neq r}\prod_{k\neq s} e^{-\Phi_{jko}v^{-\theta}}\left[\prod_{j=1}^{R}\prod_{k=1}^{S}\prod_{m\neq o} e^{-\Phi_{jkm}v^{-\theta}}\right]\theta\Phi_{rso}v^{-(\theta+1)}e^{-\Phi_{rso}v^{-\theta}}dv, \\
=& \int_0^\infty \prod_{j=1}^{R}\prod_{k=1}^{S}\prod_{m=1}^{O} \theta\Phi_{rso}v^{-(\theta+1)}e^{-\Phi_{jkm}v^{-\theta}}dv, \\
=& \int_0^\infty \theta\Phi_{rso}v^{-(\theta+1)}e^{-\Phi}v^{-\theta}dv.
\end{aligned}$$

Note that:

$$\frac{d}{dv}\left[-\frac{1}{\Phi}e^{-\Phi v^{-\theta}}\right] = \theta v^{-(\theta+1)}e^{-\Phi}.$$

Therefore:

$$\lambda_{rso} = \left[-\frac{\Phi_{rso}}{\Phi}e^{-\Phi v^{-\theta}}\right]_0^\infty,$$

which becomes:

$$\lambda_{rso} = \frac{\Phi_{rso}}{\Phi},$$

where

$$\frac{\Phi_{rso}}{\Phi} = \frac{\frac{T_{rso}w_{ro}^{\theta}(1-\ell_o)^{\theta/\beta}\ell_o^{\theta\phi_o}\bar{h}_o^{\theta}}{P_{Mr}^{\alpha\theta}P_{Nr}^{(1-\alpha)\theta}}}{\sum_{j=1}^{R}\sum_{k=1}^{S}\sum_{m=1}^{O}\frac{T_{jkm}w_{jm}^{\theta}(1-\ell_m)^{\theta/\beta}\ell_m^{\theta\phi_m}\bar{h}_m^{\theta}}{P_{jM}^{\alpha\theta}P_{jN}^{(1-\alpha)\theta}}}.$$

Summing across sectors s, the probability that a worker chooses region r and occupation o is:

$$\lambda_{ro} = \sum_{s=1}^{S}\lambda_{rso} = \frac{\Phi_{ro}}{\Phi}, \qquad \Phi_{ro} = \sum_{s=1}^{S}\Phi_{rso} = \sum_{s=1}^{S}\frac{T_{rso}w_{ro}^{\theta}\left(1-\ell_o\right)^{\theta/\beta}\ell_o^{\theta\phi_o}\bar{h}_o^{\theta}}{P_{Mr}^{\theta\alpha}P_{Nr}^{\theta(1-\alpha)}}.$$

We next derive average earnings ($\overline{wage}_{ro}$). Average earnings in a person's chosen region and occupation depends on the wage per effective unit of labor, her human capital accumulation and her average ability conditional on choosing that region and occupation:

$$\overline{wage}_{ro} = \mathbb{E}\left[w_{ro}h_o z \mid \text{chooses } r \text{ and } o\right]. \tag{A91}$$

To derive the distribution of ability conditional on choosing a region and occupation, note that an implication of the Fréchet distribution is that the distribution of utility conditional on choosing a region, sector and occupation is the same for all regions, sectors and occupations, and is equal to the distribution of utility across all regions, sectors and occupations (A88), reproduced below:

$$F\left(v\right) = e^{-\Phi v^{-\theta}}, \qquad \Phi = \sum_{j=1}^{R}\sum_{k=1}^{S}\sum_{m=1}^{O}T_{jkm}\bar{w}_{jm}^{\theta}.$$

Therefore the distribution of utility conditional on choosing a region and occupation is the same as the distribution of utility conditional on choosing a region, sector and occupation. Using (A88) and the monotonic relationship between utility and ability:

$$v = \bar{w}_{ro} z,$$

we obtain the distribution of ability conditional on choosing a region and occupation :

$$F(z) = e^{-\Phi^* z^{-\theta}}, \qquad \Phi^* = \sum_{j=1}^{R} \sum_{k=1}^{S} \sum_{m=1}^{O} T_{jkm} \left(\bar{w}_{jm} / \bar{w}_{jo} \right)^{\theta}. \tag{A92}$$

Therefore expected ability in a worker's chosen region and occupation is:

$$\mathbb{E}\left[z \mid \text{chooses } r \text{ and } o \right] = \int_0^\infty \theta \Phi^* z^{-\theta} e^{-\Phi^* z^{\theta}} dz.$$

Defining the following change of variables:

$$y = \Phi^* z^{-\theta}, \qquad dy = \theta \Phi^* z^{-(\theta+1)} dz,$$

expected ability can be re-written as:

$$\mathbb{E}\left[z \mid \text{chooses } r \text{ and } o \right] = \int_0^\infty y^{-1/\theta} \left(\Phi^* \right)^{1/\theta} e^{-y} dy,$$

which yields:

$$\mathbb{E}\left[z \mid \text{chooses } r \text{ and } o \right] = \gamma \left(\Phi^* \right)^{1/\theta}, \qquad \gamma = \Gamma \left(\frac{\theta - 1}{\theta} \right), \tag{A93}$$

where $\Gamma(\cdot)$ is the gamma function. Now note that:

$$\lambda_{ro} = \frac{\sum_k T_{rko} \bar{w}_o^{\theta}}{\sum_{k=1}^{S} \sum_{m=1}^{O} T_{rkm} \bar{w}_{rm}^{\theta}}, \tag{A94}$$

which can be written as:

$$\lambda_{ro} = \frac{\sum_k T_{rko}}{\sum_{k=1}^{S} \sum_{m=1}^{O} T_{rkm} \left(\bar{w}_{rm} / w_{ro} \right)^{\theta}} = \frac{\sum_k T_{rko}}{\Phi^*}.$$

Therefore expected ability in a worker's chosen region and occupation (A93) can be re-expressed as:

$$\mathbb{E}\left[z \mid \text{chooses } r \text{ and } o \right] = \gamma \left(\frac{\sum_{k=1}^{S} T_{rko}}{\lambda_{ro}} \right)^{1/\theta}, \qquad \gamma = \Gamma \left(\frac{\theta - 1}{\theta} \right). \tag{A95}$$

It follows that average human capital in a worker's chosen region and occupation is:

$$\mathbb{E}\left[h_o z \mid \text{chooses } r \text{ and } o \right] = \gamma \bar{h}_o \ell_o^{\phi_o} \left(\frac{\sum_{k=1}^{S} T_{rko}}{\lambda_{ro}} \right)^{1/\theta}. \tag{A96}$$

Using the choice probability (A94), average human capital in a worker's chosen region and occupation can be re-written as:

$$\mathbb{E}\left[h_o z \mid \text{chooses } r \text{ and } o \right] = \gamma \bar{h}_o \ell_o^{\phi_o} \frac{1}{\bar{w}_{ro}} \left(\sum_{k=1}^{S} \sum_{m=1}^{O} T_{rkm} \bar{w}_{rm}^{\theta} \right)^{1/\theta}. \tag{A97}$$

Using the definition of $\bar{w}_{ro}$ in (A86), it follows that average earnings in a worker's chosen region and occupation are:

$$\mathbb{E}\left[w_{ro}h_o z \mid \text{chooses } r \text{ and } o\right] = \gamma \left(1-\ell_o\right)^{-1/\beta} P_{Mr}^{\alpha} P_{Nr}^{1-\alpha} \left(\sum_{k=1}^{S}\sum_{m=1}^{O} T_{rkm}\bar{w}_{rm}^{\theta}\right)^{1/\theta}. \tag{A98}$$

Therefore relative average earnings only varies across occupations within regions because of variation in human capital investments ℓ_o:

$$\frac{\mathbb{E}\left[w_{ro}h_o z \mid \text{chooses } r \text{ and } o\right]}{\mathbb{E}\left[w_{rm}h_m z \mid \text{chooses } r \text{ and } m\right]} = \left(\frac{1-\ell_o}{1-\ell_m}\right)^{-1/\beta}. \tag{A99}$$

■

Therefore this extension to multiple regions again preserves the key properties of the baseline model. Differences across regions in the average effectiveness of performing tasks within each sector and occupation (T_{rso}) generate variation across regions in employment shares (λ_{rso}) in each sector and occupation. In contrast, the return to human capital investments in each occupation (ϕ_o) affects both employment shares (λ_{rso}) and average earnings ($\overline{wage}_{ro}$). One difference in this extension to multiple regions is that the tradeables and non-tradeables consumption prices (P_{Mr}, P_{Hr}) no longer cancel from the choice probabilities in Proposition 4, because these prices vary across regions. Nonetheless, relative average earnings ($\overline{wage}_{ro}$) across occupations within the same region only depend on relative human capital investments (ℓ_o), because goods and housing prices $\{P_{Mr}, P_{Nr}\}$ are the same across occupations within regions.

A5 Data

The data section of the paper provides an overview of the data sources and definitions. This section of the web appendix provides further detail. Our main data source on employment and worker characteristics is the individual-level records from Integrated Public Use Microdata Series (IPUMS) from 1880-2000: see Ruggles, Alexander, Genadek, Goeken, Schroeder, and Sobek (2010). We construct our datasets to make maximum use of available sample sizes, wage and education data, and geographic identifiers. We use the 100 percent samples for 1880 and 1940 and the largest available sample size for all other years (typically 5 percent).[13] Wages and education are reported from 1940 onwards. We define the hourly wage as total pre-tax wage and salary income divided by annual hours worked. We adjust wage and salary income for top-coding, by multiplying top-coded values for each year by 1.5, following Autor, Katz, and Kearney (2008). We calculate annual hours worked as annual weeks worked multiplied by weekly hours worked. For annual weeks worked, we use the intervalled variable, which is available for 1940, 1960, 1980, and 2000. We take the middle of the interval values reported. For weekly hours, no single variable is available for all years. Therefore we take usual hours worked for 2000 and hours worked last week for 1940, 1960 and

[13]In robustness tests, we also report some results using the 1860 data, in which the number of occupations reported is substantially smaller than after 1880.

1980. For the year 1960, we use the intervalled variable, taking the mid-value of the intervals and using 60 hours per week for the top weekly hours category. We also use the estimates of wages by occupation in 1880 from Preston and Haines (1991), as used in Abramitzky, Boustan, and Eriksson (2012, 2014).

In the Mincer regressions in Sections 5 and 7 of the paper and in the linear probability model in Section 7 of the paper, we include the following indicator variables for observed worker characteristics: age in years; gender (male/female); educational attainment (none, nursery school to grade 4, grades 5-8, grade 9, grade 10, grade 11, grade 12, 1 year of college, 2 years of college, 3 years of college, 4 years of college, or 5+ years of college); and ethnicity (white, black, American Indian, Chinese, Japanese, or other Asian or Pacific Islander). In the linear probability model in Section 7 of the paper, we include as an additional control an indicator variable for occupations typically undertaken in headquarters. To be conservative, we consider a broad definition of headquarters occupations (IPUMS 1950 occupation codes in parentheses): Accountants and auditors (0); College Presidents and Deans, Professors and Instructors (10); Lawyers and Judges (55); Buyers and Department Heads, Store (200); Buyers and Shippers, Farm Products (201); Credit Men (204); Floormen and Floor Managers, Store (205); Inspectors, Public Administration (210); Managers and Superintendents, Building (230); Officials and Administrators (n.e.c.), Public Administration (250); Officials, Lodge, Society, Union, etc (260); Purchasing Agents and Buyers (n.e.c.) (280); Managers, officials, and proprietors (n.e.c.) (290); Agents (n.e.c.) (300); Bookkeepers (310); Messengers and Office Boys (340); Office Machine Operators (341); Shipping and Receiving Clerks (342); Stenographers, Typists, and Secretaries (350); Telegraph Operators (365); Telephone Operators (370); and Clerical and kindred workers (n.e.c.) (390).

A6 Trends in Task Inputs 1880-2000

In this section, we report additional empirical results that are discussed in Section 5 of the paper.

A6.1 Evidence from Numerical Scores

Figure 1 in the paper shows the employment share-weighted average of each numerical score (non-routine interactive, non-routine analytic, routine analytic, routine cognitive and routine non-manual) over time. Numerical scores are expressed as percentiles and weighted by occupational employment shares in each year. Each series is expressed as an index relative to its value in 1880 (so that each series takes the value one in 1880). We find a rise in inputs of non-routine interactive and non-routine analytic tasks in the late-nineteenth and early-twentieth centuries following an earlier revolution in information and communication technology, centered on the typewriter and telephone.

In Figure A1, we show that this pattern is not apparent in the preceding 1860-1880 period. We display the employment share-weighted average of each numerical score (non-routine interactive, non-routine analytic, routine analytic, routine cognitive and routine non-manual) over this preceding twenty-year period. Numerical scores are expressed as percentiles and weighted by occupational employment shares in each year. Each series is expressed as an index relative to its value in 1860 (so that each series takes the value one in 1860). As apparent from the figure, inputs of non-routine interactive and non-routine analytic tasks actually decline during this preceding twenty-year period. We report these results as a robustness test rather than as part of our main specification, because the number of occupations reported in the data is substantially smaller for years before 1880.

To tighten the link between the rise in inputs of non-routine tasks and the earlier revolution in information and communication technology, Figure A2 displays data on these technologies from the Historical Statistics of the United States Millennial Edition. As evident from the figure, the late-nineteenth and early-twentieth centuries were the period of the most rapid dissemination of these earlier information and communication technologies. In 1876, Alexander Graham Bell was the first to be granted a United States patent for the telephone. As shown in Panel A, the number of telephones grows by a factor of more than 250 from 47,900 in 1880 to 13,411,400 in 1920. In 1835, the electrical telegraph was developed and patented in the United States by Samuel Morse. As shown in Panel B, the number of miles of telegraph wire more than quadruples from 291,000 in 1880 (the first year for which data are reported) to 1,318,000 in 1902. Telegraph miles continue to increase more gradually until 1943, before declining thereafter as the telegraph is superseded by other communication technologies (in particular the telephone). The first typewriter to be commercially successful was invented in 1868 by Christopher Latham Sholes, Carlos Glidden and Samuel W. Soule in Milwaukee, Wisconsin. As shown in Panel C, physical output of typewriters more than trebles from 145,000 in 1900 (the first year for which data are reported) to 489,000 in 1921. Typewriter output continues to increase with large fluctuations to a peak of 1,925,000 in 1979, before declining sharply thereafter as typewriters are made obsolete by the computer.

A6.2 Evidence from Individual Production Tasks

No additional empirical results.

A6.3 Tasks and Wages

Quantification Results We now report the results of the quantification of the model discussed in Sections 2.5 and 5.3 of the paper and in section A2.5 of this web appendix. We follow Hsieh, Hurst, Jones, and Klenow (2013) in assuming a value for the Fréchet shape parameter determining worker comparative advantage of $\theta = 3.44$ and a value for the parameter governing the tradeoff between consumption and human capital accumulation of $\beta = 0.693$. Using these assumed parameters and observed average earnings for each occupation ($\overline{wage}_o$), we solve for the unobserved return to human capital accumulation (ϕ_o) from the following system of equations for each occupation o:

$$\frac{\overline{wage}_o}{\prod_{o=1}^{O} \left[\overline{wage}_o\right]^{\frac{1}{O}}} = \frac{(1-\ell_o)^{-1/\beta}}{\prod_{o=1}^{O} \left[(1-\ell_o)^{-1/\beta}\right]^{\frac{1}{O}}} = \frac{(1+\beta\phi_o)^{-1/\beta}}{\prod_{o=1}^{O} \left[(1+\beta\phi_o)^{-1/\beta}\right]^{\frac{1}{O}}}. \tag{A100}$$

Given the resulting solutions for ϕ_o, we compute human capital investments $\ell_o = 1/(1 + 1/\beta\phi_o)$, and construct a measure of the contribution of human capital investments to observed employment shares ($\mathbb{B}_{so} = (1-\ell_o)^{\theta/\beta}\,\ell_o^{\theta\phi_o}$). Using this measure and observed employment shares (λ_{so}), we solve for an adjusted measure of task effectiveness for each sector and occupation ($\mathbb{A}_{so}$):

$$\frac{\mathbb{A}_{so}}{\left[\prod_{s=1}^{S}\prod_{o=1}^{O}\mathbb{A}_{so}\right]^{\frac{1}{O}}} = \frac{\lambda_{so}/\left[\prod_{s=1}^{S}\prod_{o=1}^{O}\lambda_{so}\right]^{\frac{1}{O}}}{\mathbb{B}_{so}/\left[\prod_{s=1}^{S}\prod_{o=1}^{O}\mathbb{B}_{so}\right]^{\frac{1}{O}}}, \tag{A101}$$

where $\mathbb{A}_{so} = T_{so}w_o^{\theta}\bar{h}_o^{\theta}$ captures variables that enter employment shares (λ_{so}) isomorphically to the Fréchet scale parameter that determines the average effectiveness of workers in performing tasks in each sector and occupation (T_{so}). As employment shares are homogeneous of degree zero in task effectiveness, task effectiveness is only uniquely identified up to a normalization. We choose the convenient normalization that the geometric mean of task effectiveness is equal to one across sectors and occupations and compare relative levels of task effectiveness across sectors and occupations.

In Figure A3, we display the cumulative distribution of human capital returns (ϕ_o) across percentiles of the occupation task distribution.[14] As the human capital return distributions are cumulative, they necessarily add up to one for each year. Furthermore, the slope of each cumulative distribution corresponds to the human capital return at that percentile of the task distribution. Occupations are sorted along the horizontal axis by non-routine percentiles in the top-left panel, by routine percentiles in the top-right panel and by manual percentiles in the bottom left panel. Looking across these three panels, we find the largest changes in the relative return to human capital investments for occupations with different levels of non-routine tasks (top left panel). Comparing 1880 and 1940 in this top left panel, the solid gray line (for 1940) typically lies below the solid black line (for 1880), with the gap largest for intermediate-high levels of non-routine tasks. Therefore the solid gray line must have a steeper slope than the solid black line for the most non-routine occupations (in order for the cumulative distribution to add up to one). This pattern implies an increase in the relative return to human capital investments for the most non-routine occupations. Comparing 1940 to 2000 in the top left panel, the dashed black line (for 2000) generally lies below the solid gray line (for 1940), with the gap largest for intermediate levels of non-routine tasks. As a result, the slope of the dashed black line is steepest relative to that of the solid gray line at intermediate-high levels of non-routine tasks. This pattern implies an increase in the relative return to human capital investments for occupations with intermediate-high levels of non-routine tasks. Therefore, taking the two sub-periods together, there is a secular increase in the relative return to human capital investments for non-routine occupations, which at first is greatest for the most non-routine occupations, before later becoming greatest for those occupations with intermediate-high levels of non-routine tasks.

[14]Specifically, each occupation has a human capital return and a task measure (percentile score) in a given year. We sort occupations by their task percentile in a given year (horizontal axis). We then cumulate the human capital returns across these percentiles of the occupation task distribution, scaling by the sum of the human capital returns to obtain a cumulative distribution (vertical axis).

In Figure A4, we show the cumulative distribution of relative task effectiveness ($\mathbb{A}_o$) across percentiles of the occupation task distribution, where task effectiveness for each occupation is measured as the employment-weighted sum across sectors within that occupation ($\mathbb{A}_o = \sum_{s=1}^{S} (E_{so}/E_o) \mathbb{A}_{so}$).[15] Comparing Figures A4 and A3, we typically find larger changes in relative task effectiveness than in relative human capital returns across occupations. Again these changes are most apparent across occupations with different levels of non-routine tasks (top left panel), although now we find substantial changes in relative task effectiveness across occupations with different levels of routine tasks (top right panel) or manual tasks (bottom left panel). Comparing 1880 and 1940 for non-routine tasks in the top left panel, the solid gray line (for 1940) is above the solid black line (for 1880) at low levels of non-routine tasks, but is below the solid black line at high levels of non-routine tasks, which implies that the solid gray line must have a steeper slope than the solid black line for both the least and the most non-routine occupations (since both lines start at zero and add up to one). This pattern implies a polarization of task effectiveness – an increase in relative task effectiveness for both the least and most non-routine occupations – in the first half of our sample period. In contrast, comparing 1940 and 2000 in the top left panel, the dashed black line (for 2000) lies almost everywhere below the solid gray line (for 1940), which implies that the dashed black line has a steeper slope than the solid gray line for more non-routine occupations (in order for the cumulative distribution to add up to one). This pattern implies a secular rise in the relative task effectiveness of more non-routine occupations in the second half of our sample period, consistent with computers complementing workers performing non-routine tasks during this period.

Comparing 1880 and 1940 for routine tasks in the top right panel, the solid gray line (for 1940) lies relatively close to the solid black line (for 1880) throughout the entire distribution. This similarity of the two distributions implies relatively small changes in relative task effectiveness across occupations with different levels of routine tasks in the first half of our sample period. In contrast, comparing 1940 and 2000 in the top right panel, the dashed black line (for 2000) has a similar slope to the solid gray line (for 1940) for the least-routine occupations, but then generically has a lower slope than the solid gray line throughout most of the rest of the distribution (with a couple of exceptions of step increases around the median and at the top of the distribution). This pattern of results implies a secular fall in the relative task effectiveness of workers performing more routine tasks relative to those performing the least routine tasks in the second half of our sample period, consistent with computers substituting for routine tasks during this period.

Comparing 1880 and 1940 for manual tasks in the bottom left panel, the solid gray line (for 1940) is below the solid black line (for 1880) for the least manual tasks, before rising to follow the solid black line, and then falling below the solid black line for most of the distribution above the median. Therefore the solid gray line has a lower slope than the solid black line for the least manual tasks and a steeper slope than the solid black line for the most manual tasks. This pattern of results implies an increase in relative task effectiveness for workers performing more manual tasks in the first half of our sample period,

[15] We construct the measures on the horizontal and vertical axes in Figure A4 in the same way as in Figure A3, except that we use the cumulative distribution of average occupation effectiveness ($\mathbb{A}_o$) on the vertical axis instead of the cumulative distribution of the occupation human capital returns (ϕ_o).

consistent with the idea that improvements in technology in the late-nineteenth and early-twentieth centuries complemented workers performing more manual tasks. In contrast, comparing 1940 and 2000 in the bottom left panel, the dashed black line (for 2000) has a steeper slope than the solid gray line (for 1940) for the least manual tasks and has a shallower slope than the solid gray line (for 1940) through much of the rest of the distribution, with a couple of exceptions in the form of step increases around the 60th and 80th percentiles. This pattern of results implies a decline in the relative task effectiveness of workers performing more manual tasks in the second half of our sample period, consistent with technological change substituting for manual tasks during this period.

Residual Wage Inequality Results In Figures A5 and Figure A6, we show that the results for overall wage inequality in Figures 4 and 6 in the paper also hold for residual wage inequality. Following Juhn, Murphy, and Pierce (1993) and Autor, Katz, and Kearney (2008), we use our individual-level Census data on annual wages (available from 1940 onwards) to estimate a Mincer regression controlling for education, age, gender and ethnicity:

$$\ln w_{it} = X_{it}\nu_t + u_{it} \tag{A102}$$

where w_{it} is the annual wage; X_{it} are observable worker characteristics (education, gender, age and ethnicity); ν_t are coefficients that we allow to differ across years to reflect changes in the premia to these observed characteristics; and u_{it} is a stochastic error. We use the estimated residuals ($\hat{u}_{it}$) from this Mincer regression as our measure of residual wage inequality controlling for worker observables. We compute the occupation residual wage as the average of the residual wages across all workers within that occupation.

In Figure A5, we display the cumulative distribution of employment shares across percentiles of the occupation residual wage distribution. On the horizontal axis of the figure, occupations are sorted in a given year according to their percentiles of the occupation residual wage distribution in that year. On the vertical axis, we display the cumulative sum of employment shares in that given year across the percentiles of the occupation wage distribution in that year. As the employment share distributions are cumulative, they necessarily add up to one for each year. Furthermore, the slope of each cumulative distribution corresponds to employment at that percentile of the wage distribution.

Figure A5 shows that we find the same pattern of rising inequality from 1940-2000 for residual wages as for actual wages (Figure 4 in the paper). At low residual wages, the distributions for the two years track one another relatively closely, but there is a smaller mass of workers at intermediate residual wages in 2000 than in 1940 (the gray line has a steeper slope at intermediate wages than the black dashed line), and a greater mass of workers at high residual wages in 2000 than in 1940 (the black dashed line has a steeper slope than the gray line at high wages). Therefore the second half of our sample period is characterized by increased residual wage inequality across occupations and a polarization of wages towards the top of residual wage distribution at the expense of the middle of the residual wage distribution.

Figure A6 shows the cumulative distribution of tasks (non-routine, routine and manual) across percentiles of the occupation residual wage distribution in 1940 and 2000.[16] The left panel shows these distributions for 1940 and the right panel shows them for 2000. Consistent with the results for overall wages in Figure 6 in the paper, we find a sharp increase in the extent to which non-routine tasks are concentrated at higher percentiles of the wage distribution from 1940-2000.

A7 Tasks and Technology

In this section, we report additional results on tasks and technology from Section 6 of the paper. We estimate the following regression using observations on sectors s and years t from 1880-2000 for the within-industry relationship between inputs of task k and use of technology m:

$$\mathbb{T}_{skt} = \beta_{tkm} \left(\mathbb{S}_{sm} \times \mathbb{I}_t \right) + \eta_s + d_t + u_{st} \tag{A103}$$

where $\mathbb{T}_{skt}$ is input of task k in sector s at time t; $\mathbb{S}_{sm}$ is the share of sector s's inputs that originate from industries that produce technology m; $\mathbb{I}_t$ is an indicator variable for year t; η_s are sector fixed effects; d_t are year dummies; and u_{st} is a stochastic error. We cluster the standard errors by sector to allow for serial correlation in the error term over time. We report standardized beta coefficients (scaled by variable standard deviations) for comparability across the different task and technology measures.

In the paper, we show that industries that use office and computing machines intensively experience an increase in inputs of non-routine tasks relative to other industries in the late-nineteenth century. This timing aligns closely with the period of rapid dissemination of a cluster of information and communication technologies (ICTs) centered on the typewriter and telephone (Figure A2). We now show that this differential trend in inputs of non-routine tasks in these industries is not present in the earlier 1860-1880 period. We augment the regression sample with this earlier twenty-year period and re-estimate equation (25) in the paper using 1860 as the excluded year. We report these results as a robustness test rather than as our main specification, because the number of occupations reported in the data is substantially smaller for years before 1880. As shown in Figure A7, we find that the estimated coefficient on office and computing machinery is flat from 1860-1880 before increasing from 1880-1900, consistent with the timing of this late-nineteenth century ICT revolution.

A8 Tasks and Cities

In this section, we report additional results on tasks and cities from Section 7 of the paper.

A8.1 Evidence from Numerical Scores

To provide further evidence on changes in task inputs in metro and non-metro areas over time, Section 7.1 of the paper uses our individual-level Census data on education (available from 1940 onwards) to estimate

[16] Specifically, each occupation has a task measure (percentile score) and a residual wage percentile in a given year. We sort occupations by their residual wage percentile in a given year. We then cumulate the task measure across these percentiles of the occupation residual wage distribution, scaling by the sum of the task measure (to obtain a cumulative distribution).

the following linear probability model for the probability that individual i is in a metro area in year t:

$$\mathbb{I}_{it}^{M} = X_{it}\iota_t + \mathbb{T}_{o(i)t}\varsigma_t + \eta_{s(i)} + \chi_{it}, \tag{A104}$$

where $\mathbb{I}_{it}^{M}$ is an indicator variable that is equal to one if individual i is in a metro area; X_{it} are observable worker characteristics (education, gender, age and ethnicity); ι_t are coefficients that we allow to differ across years to reflect changes in the premia to these observed characteristics; $\mathbb{T}_{o(i)t}$ are the numerical score measures of the tasks undertaken by worker i within her occupation $o(i)$ at time t; ς_t are task premia that we again allow to change over time; $\eta_{s(i)}$ is a fixed effect for worker i's sector $s(i)$; and χ_{it} is a stochastic error. Although we focus on a linear probability model to facilitate the inclusion of sector fixed effects, we find a similar pattern of results using a Probit specification.

We discuss the results of this specification in Section 7.1 of the paper. In Table A1 of this web appendix, we report the estimated coefficients and standard errors for (A104) for each year separately from 1940 onwards. Column (1) of Table A1 estimates (A104) including our controls for observable worker characteristics (education, gender, age and ethnicity) and the three task measures (non-routine, routine and manual). We find an increase in the estimated non-routine coefficient and a decline in the estimated routine coefficient over time, which implies that the observed changes in task inputs cannot be fully explained by changes in educational attainment or demographic composition. Column (2) of Table A1 augments the specification from Column (1) with a full set of industry fixed effects. We continue to find an increased concentration of non-routine tasks in metro areas and a reduced concentration of routine tasks in metro areas over time, confirming that our findings are not driven by a change in industry composition. Column (3) of Table A1 further augments the specification from Column (2) with an indicator variable for occupations typically undertaken in headquarters.[17] Even in this specification including the full set of controls, we continue to find a similar pattern of results. Whereas non-routine tasks were statistically significantly *less* likely to be performed in metro areas in 1940, they were statistically significantly *more* likely to be undertaken in metro areas in 2000, confirming that we find a reversal in the pattern of task specialization in urban and rural areas over time, even after controlling for observed worker characteristics and sectoral and functional specialization.

A8.2 Evidence from Individual Production Tasks

In Section 7.2 of the paper, we provide finer resolution evidence on the change in task inputs in urban areas relative to rural areas, using our new methodology for measuring individual tasks. We examine which verbs are most concentrated in metro areas by regressing the share of employment that is located in metro areas within a sector and occupation on the frequency with which a verb is used for that occupation. In particular, for each verb v and year t from 1880-2000, we estimate the following regression using observations across occupations o and sectors s for a given verb and year:

$$\text{MetroShare}_{ost} = \alpha_{vt}\text{VerbFreq}_{vo} + \eta_{vst} + \varepsilon_{ost}, \tag{A105}$$

[17] See Section A5 of this web appendix for the list of occupations typically undertaken in headquarters.

where MetroShare_{ost} is the share of employment within occupation o and sector s that is located in metro areas in year t; VerbFreq_{vo} is defined in Section 4 of the paper for verb v and occupation o; η_{vst} are sector fixed effects for verb v and year t; and ε_{ost} is a stochastic error.

In Section 7.2 of the paper, we report the results of estimating (A105) using the verbs from 1991 occupational descriptions. In Table A2, we reports the results of estimating this specification using the verbs from the 1939 occupational descriptions. We find a similar pattern of changes in task inputs in metro areas relative to non-metro areas over time. The verbs most correlated with metro employment shares in 1880 include physical tasks such as "Slot," "Thread," "Stitch" and "Straighten." In contrast, the verbs most correlated with metro employment shares in 2000 include analytical and interactive tasks such as "Advise," "Audit," "Question" and "Present." Therefore, whether we use 1939 or 1991 occupational descriptions, we find a reversal in the types of individual production tasks most concentrated in urban areas, highlighting a transformation in the nature of agglomeration. While the typical urbanite in 1880 was likely to be employed in a manual task rearranging the physical world, their counterpart in 1940 was most frequently engaged in recording and processing information, and the modern city dweller typically performs tasks involving ideas, initiative and interaction.

A8.3 Tasks and Wage Inequality

In Figure A8, we show that the results for overall wage inequality in Figure 12 in the paper also hold for residual wage inequality. Following Juhn, Murphy, and Pierce (1993) and Autor, Katz, and Kearney (2008), we use our individual-level Census data on annual wages (available from 1940 onwards) to estimate a Mincer regression controlling for education, age, gender and ethnicity, as discussed in the paper. We use the estimated residuals from this Mincer regression as our measure of residual wage inequality controlling for worker observables. We compute the occupation residual wage for metro areas as the average of the residual wages across all workers within that occupation in metro areas. We compute the occupation residual wage for non-metro areas analogously.

In Figure A8, we display the cumulative distribution of employment shares across percentiles of the occupation residual wage distribution in each year for metro and non-metro areas separately. On the horizontal axis of the figure, occupations are sorted in a given year according to their percentiles of the occupation residual wage distribution for metro or non-metro areas in that given year. On the vertical axis, we display the cumulative sum of employment shares for metro or non-metro areas in that given year across these percentiles of the residual wage distribution. As the employment share distributions are cumulative, they necessarily add up to one for each year. Furthermore, the slope of each cumulative distribution corresponds to employment at that percentile of the wage distribution.

As shown in Figure A8, we find much larger changes in the distribution of employment across percentiles of the residual wage distribution in metro areas than in non-metro areas from 1940-2000. Therefore most of the increase in residual wage inequality and the polarization of the residual wage distribution towards higher wages in Figure A5 is driven by metro areas. These results for residual wage inequality are consistent with our results for overall wage inequality in the paper.

A9 Robustness

In this section, we report the results of a number of additional robustness tests and empirical specifications that are discussed in the paper.

A9.1 Variation within Metro Areas

In Section 7 of the paper, we show that there is a reversal in the types of tasks most concentrated in metro areas relative to non-metro areas over our long historical time period. As shown in Figure 10 in the paper, in 1880, metro areas were *less* specialized in non-routine tasks than non-metro areas. In contrast, in 2000, metro areas were *more* specialized in non-routine tasks than non-metro areas.

In this section of the web appendix, we provide further evidence in support of this transformation in the organization of economic activity in more-densely-populated locations relative to less-densely-populated locations, using a different source of variation within metro areas rather than between metro and non-metro areas. In Figure A9, we display the employment-weighted average of the occupation non-routine numerical score for each metro area in 1880 and 2000. Non-routine numerical scores are expressed as percentiles and weighted by occupational employment shares in a given metro area in a given year. To ensure that metro areas correspond to meaningful economic units in each year, we use time-varying definitions of metro areas from IPUMS. Therefore both the boundaries of each metro area and the number of metro areas change over time. In 1880, we find little relationship between specialization in non-routine tasks and log population density across metro areas, which is reflected in a statistically insignificant OLS coefficient (standard error) of 0.0042 (0.0037). In contrast, in 2000, we find a strong positive and statistically significant relationship between specialization in non-routine tasks and log population density, which is reflected in an OLS coefficient (standard error) of 0.0075 (0.0011). We also find the same strong positive relationship for 2000 if we restrict the sample to 1880 metro areas.

Therefore, using a different source of variation within metro areas with different population densities, we find a similar pattern of results as between metro and non-metro areas. We find a transformation in the nature of agglomeration, with non-routine tasks increasingly concentrated in metro areas rather than non-metro areas, and increasingly concentrated in metro areas with higher population densities.

A9.2 Comparison with Other Task Measures

In the paper, we report a number of specification checks on our measures of the individual production tasks for each occupation. In this section of the web appendix, we report additional robustness tests and comparisons with other task measures.

As discussed in Section 4 of the paper, our baseline task measures use time-invariant occupational descriptions from the 1991 Dictionary of Occupational Titles (DOTs) for all years in our sample. Therefore, our results are not driven by changes in the language used in the occupational descriptions over time, because these occupational descriptions are held fixed for all years. Our use of time-invariant occupational descriptions implies that changes in aggregate task input over time are driven solely by changes in employment shares across occupations with different (time-invariant) task inputs. Although Roget's Thesaurus continues to be the seminal reference for English language, one remaining concern could be that it was compiled in a different year (1911) from the occupational descriptions (1991). In Section A8.2 of this web appendix, we report a robustness test using occupational descriptions from the first edition of the DOTs in 1939 and demonstrate a similar pattern of results. This similarity of the results using occupational descriptions more than fifty years apart provides strong evidence that the difference in dates between the thesaurus and occupational descriptions is not consequential for our results.[18]

As a further check on our findings, the paper reports results using the numerical scores summarizing job requirements that have been used in previous research for recent decades. When we aggregate our measures of individual production tasks to the level of thesaurus subdivisions, we find a consistent pattern of results with the non-routine, routine and manual numerical scores (see for example Figures 1-3 in the paper). We also use these numerical scores to provide further evidence on changes in task input within occupations over time by comparing numerical score measures from the 1991 DOTs (as used in Autor, Levy, and Murnane 2003) with those from the 1949 DOTs (as used in Gray 2013). Consistent with the results in Spitz-Oener (2006), we find some evidence of within-occupation changes in task inputs. However, all of our empirical specifications use percentile rankings of occupations. We find that these percentile rankings are strongly positively correlated between the historical and contemporary measures.[19]

As an additional check, we compared our measures of tasks based on thesaurus sections with those from the Occupational Information Network (O*NET). Whereas the numerical scores in the DOTs are based on ratings from expert job evaluations, the measures of occupation characteristics in O*NET are largely based on employee and employer surveys. In Table A3 of this web appendix, we report the correlation across occupations between a measure of interactiveness based on the frequency of verb use from Classes 4 and 5 of the thesaurus (the "Intellectual Faculties" and "Voluntary Powers" respectively) with seventeen subcategories of the work activity "Interacting with Others" from the Occupational Information Network (O*NET). These measures were constructed by US Department of Labor/Employment and Training Administration (USDOL/ETA) based on questionnaires about detailed work activities issued to a random sample of businesses and workers. Panel A reports unweighted correlations, while Panel B reports correlations weighted by employment. The measures cover a wide range of forms of interaction, including "Assisting and caring for others" and "Resolving conflict and negotiating with others."

[18] As a check on whether changes in language use over time are correlated with the nature of production tasks, we compared the frequency distribution of verbs across sections of the thesaurus using 1939 and 1991 occupational descriptions. We find that the two frequency distributions are strongly correlated across sections of the thesaurus, with no evident pattern to the differences between them.

[19] Correlating the 1949 and 1991 percentile numerical scores across the occupations in the data in 2000, we find the following correlations: non-routine (0.799), routine (0.770) and manual (0.700), suggesting a high correlation in the ranking of occupations in terms of task input over time.

We find that all of the correlations with our measure of interactiveness are positive and statistically significant. The five categories with the highest unweighted correlations correspond closely to activities involving thought, communication and intersocial activity: "Interpreting the meaning of information for others," "Provide consultation and advice to others," "Resolving conflict and negotiating with others," "Establishing and maintaining interpersonal relationships" and "Performing administrative activities." These positive correlations provide external validation that aggregations of our measures of individual tasks based on the frequency with which verbs appear in occupational descriptions and sections of the thesaurus are consistent with measures of tasks based on employee and employer surveys.

Therefore, we find a consistent pattern of results across a range of alternative measures of the tasks performed by workers within occupations – whether we use contemporary or historical occupational descriptions, whether we use occupational descriptions or numerical scores, whether we use contemporary or historical numerical scores, and whether we use measures of occupation characteristics from job evaluations or employee and employer surveys. This consistency of our findings across these quite different approaches suggests that our results are indeed capturing robust and systematic features of the data. A key advantage of our methodology based on occupational descriptions and thesaurus sections is that we are able to track production tasks at a much higher resolution than has hitherto been possible, enabling us to trace patterns of complementarity and substitutability between technologies and individual production tasks. We show in the paper that the aggregate measures of tasks used in existing research mask much larger changes in task inputs using our more disaggregated measures of tasks.

References

Abramitzky, R., L. P. Boustan, and K. Eriksson (2012): "Europe's Tired, Poor, Huddled Masses: Self-selection and Economic Outcomes in the Age of Mass Migration," *American Economic Review*, 102(5), 1832–1856.

——— (2014): "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration," *Journal of Political Economy*, 122(3), 467–506.

Ahlfeldt, G., S. Redding, D. Sturm, and N. Wolf (2015): "The Economics of Density: Evidence from the Berlin Wall," *Econometrica*, 83(6), 2127–2189.

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

Autor, D. H., F. Levy, and R. J. Murnane (2003): "The Skill Content of Recent Technological Change: An Empirical Exploration," *Quarterly Journal of Economics*, pp. 1279–1333.

Burstein, A., E. Morales, and J. Vogel (2015): "Accounting for Changes in Between-Group Inequality," *NBER Working Paper*, 20855.

Davis, M. A., and F. Ortalo-Magné (2011): "Housing Expenditures, Wages, Rents," *Review of Economic Dynamics*, 14(2), 248–261.

Eaton, J., and S. Kortum (2002): "Technology, Geography, and Trade," *Econometrica*, 70(5), 1741–1779.

Gray, R. (2013): "Taking Technology to Task: The Skill Content of Technological Change in Early Twentieth Century United States," *Explorations in Economic History*, 50(3), 351–367.

Heckman, J. J., and B. E. Honore (1990): "The Empirical Content of the Roy Model," *Econometrica*, 58(5), 1121–1149.

Hsieh, C.-T., E. Hurst, C. I. Jones, and P. J. Klenow (2013): "The Allocation of Talent and U.S. Economic Growth," *NBER Working Paper*, 18693.

Juhn, C., K. M. Murphy, and B. Pierce (1993): "Wage Inequality and the Rise in Returns to Skill," *Journal of Political Economy*, 101(3), 410–442.

Lagakos, D., and M. E. Waugh (2013): "Selection, Agriculture, and Cross-Country Productivity Differences," *American Economic Review*, 103(2), 948–80.

McFadden, D. (1974): "The Measurement of Urban Travel Demand," *Journal of Public Economics*, 3(4), 303–328.

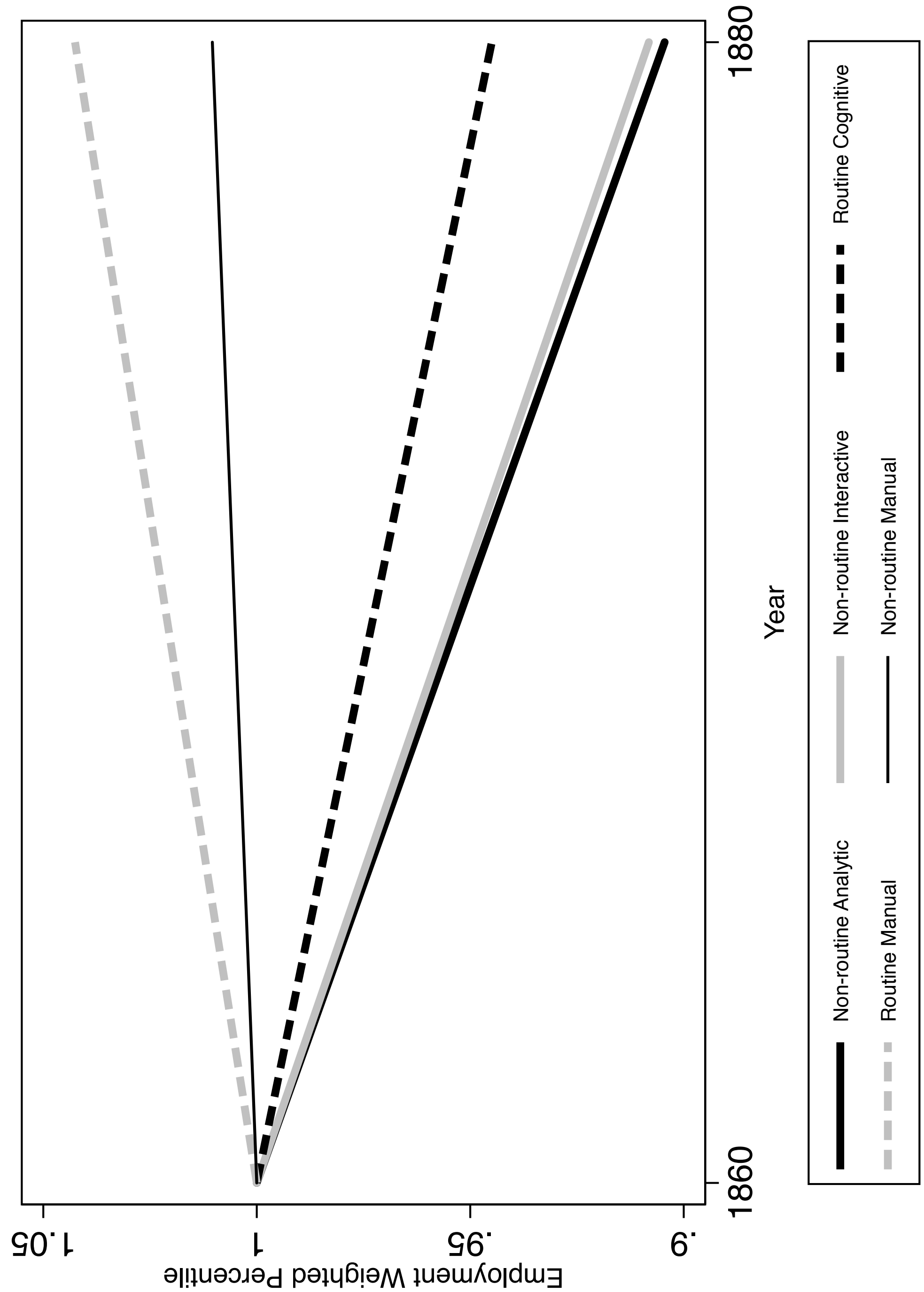
Preston, S. H., and M. R. Haines (1991): *Fatal Years: Child Mortality in Late Nineteenth-Century America.* Princeton University Press, Princeton, NJ.

Roy, R. D. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3, 135–146.

Ruggles, Steven, J., T. Alexander, K. Genadek, R. Goeken, M. B. Schroeder, and M. Sobek (2010): "Integrated Public Use Microdata Series," University of Minnesota, Minneapolis.

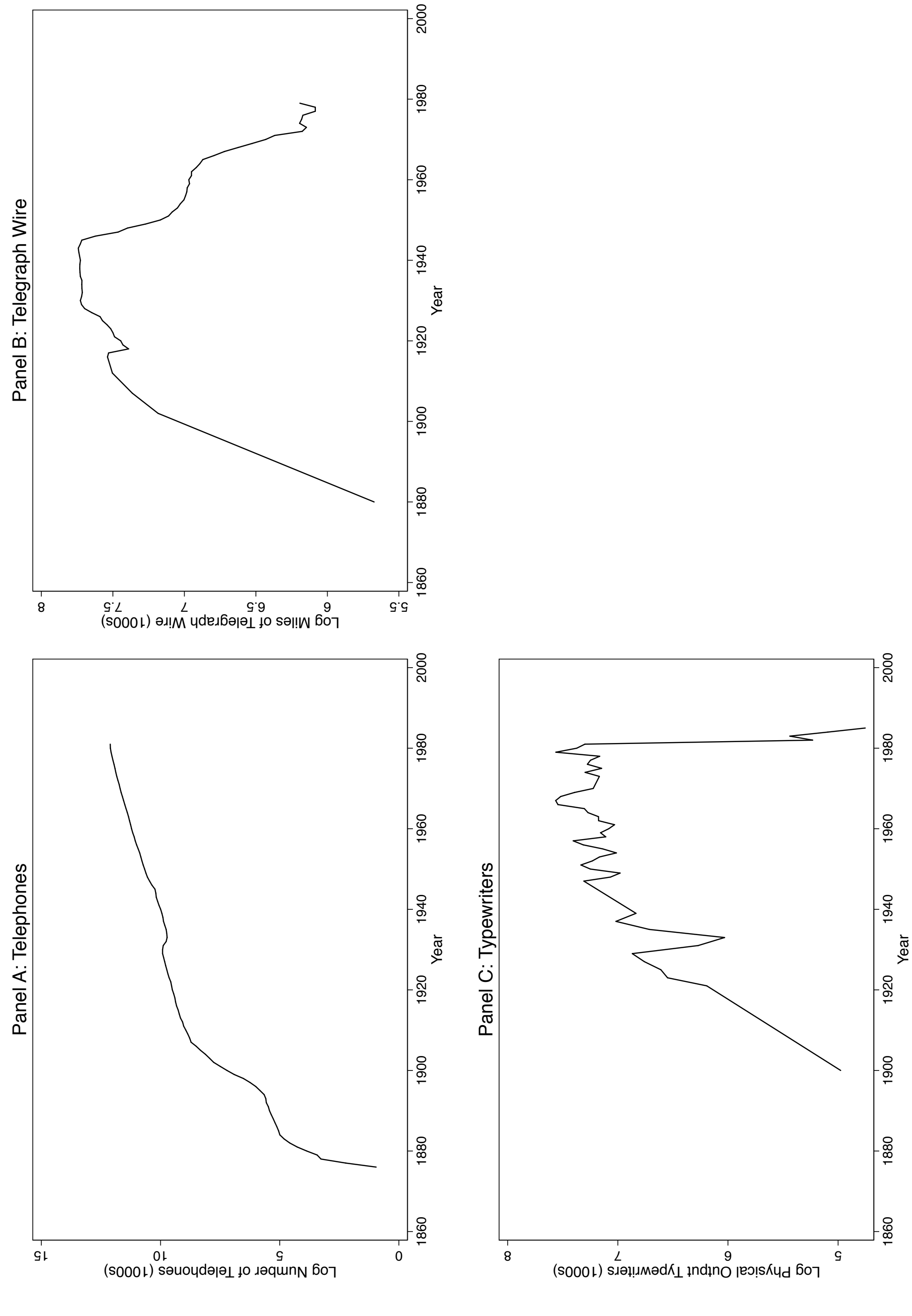
Spitz-Oener, A. (2006): "Technical Change, Job Tasks, and Rising Educational Demands: Looking outside the Wage Structure," *Journal of Labor Economics*, 24(2), 235–270.

Figure A1: Task Input by Numerical Score from 1860-1880



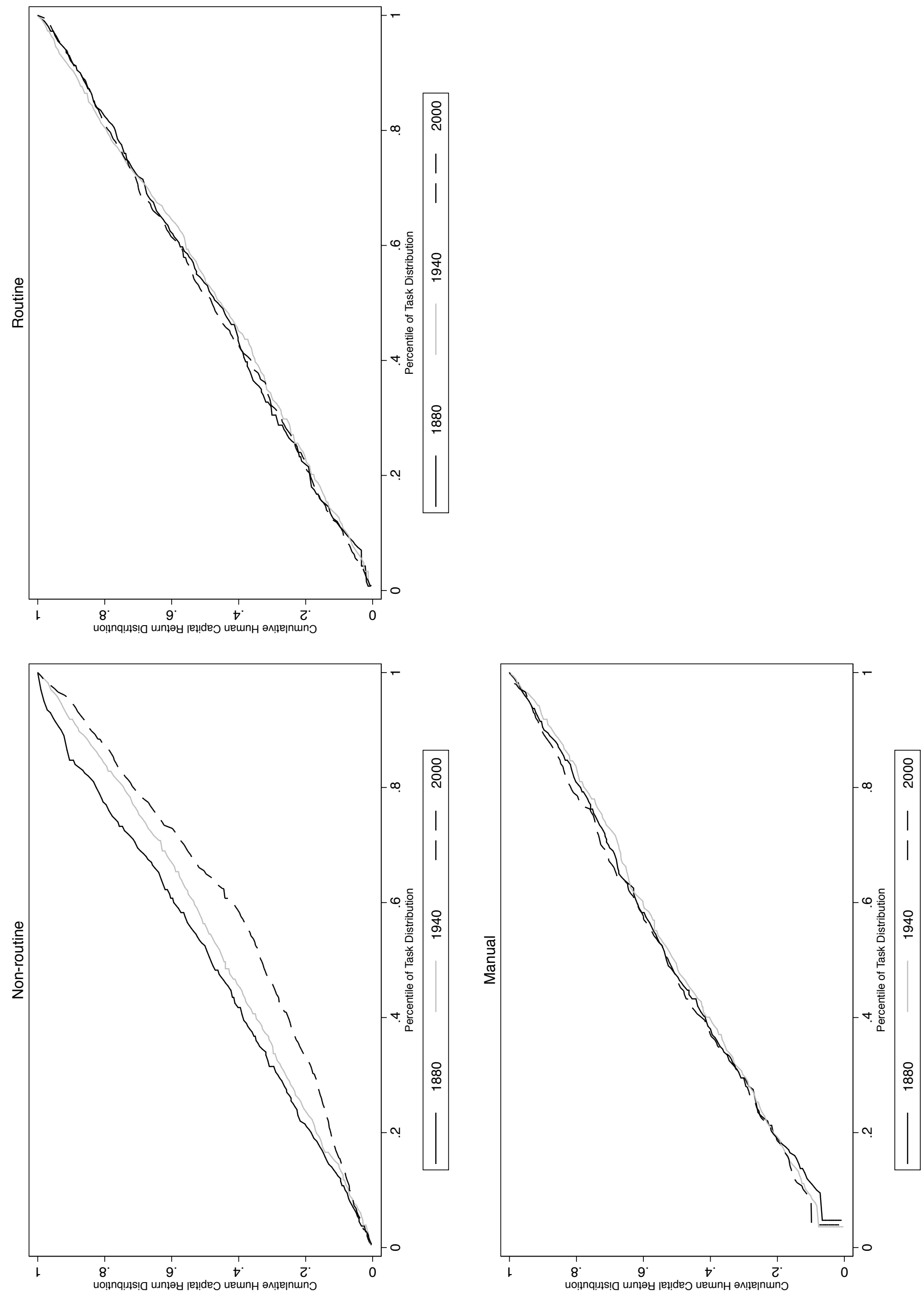
Note: Employment-weighted average of occupation numerical scores summarizing job requirements (non-routine analytic, non-routine interactive, routine cognitive, routine manual and non-routine manual) for each year, expressed as an index relative to 1860. Occupation numerical scores from the Dictionary of Occupational Titles (DOTs) for 1991. The time-invariant numerical score for each occupation is converted into the percentile of its distribution across occupations. Employment in each occupation and year measured using IPUMS population census data. This figure is constructed in the same way as Figure 1 in the paper, except that the time period is different and the index is normalized to equal one in 1860 rather than 1880.

Figure A2: Late-Nineteenth Century Improvements in Information and Communication Technology (ICT)



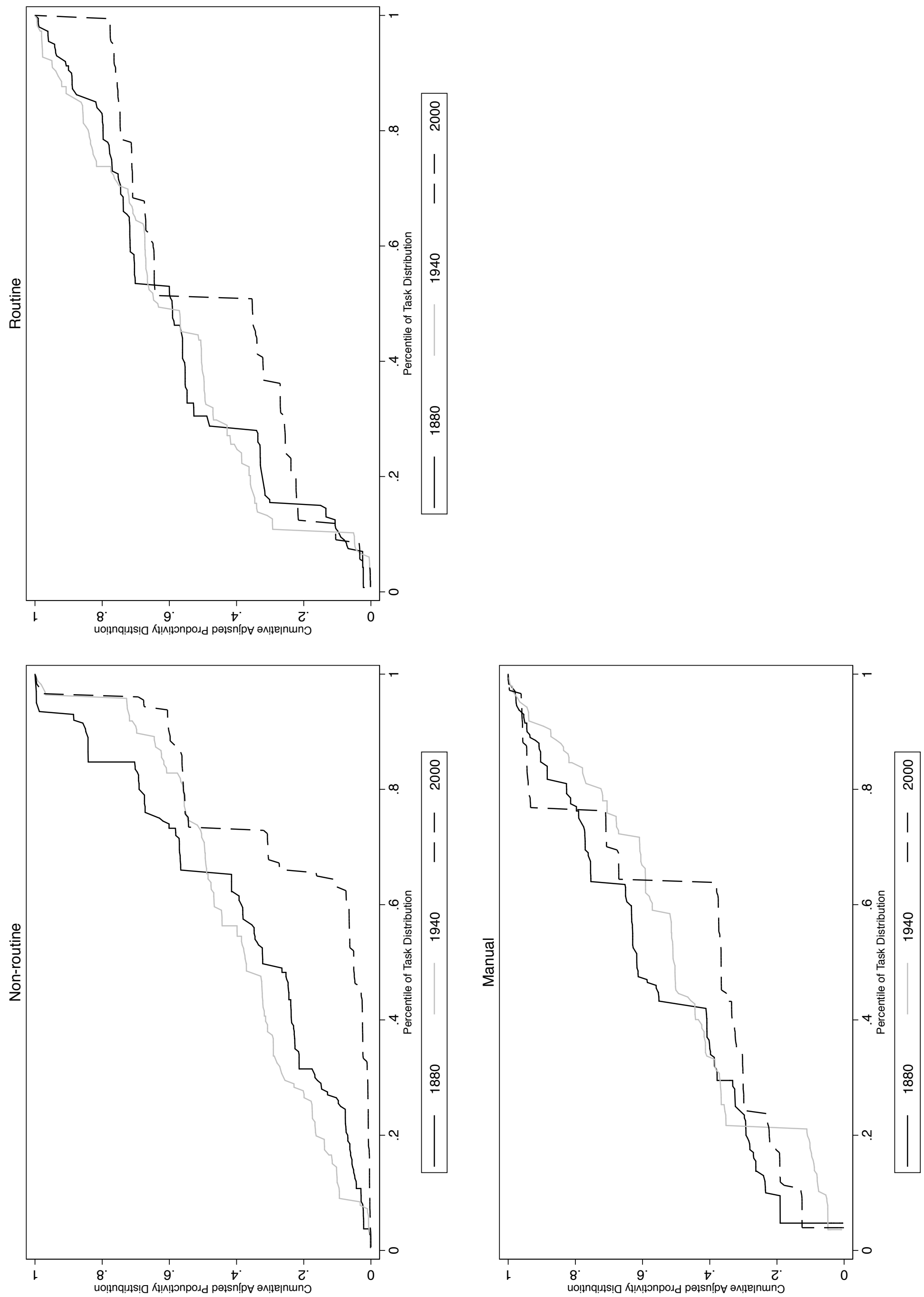
Note: Log number of telephones (1,000s), log number of miles of telegraph wire (1,000s) and log physical output of typewriters (1,000s) from Historical Statistics of the United States Millennial Edition. The first commercially-successful typewriter was invented in 1868 by Christopher L. Sholes, Carlos Glidden and Samuel W. Soule in Milwaukee, Wisconsin. In 1876, Alexander Graham Bell was the first to be granted a United States patent for the telephone. The electrical telegraph was developed and patented in the United States by Samuel Morse somewhat earlier in 1837.

Figure A3: Cumulative Distribution of Human Capital Returns (ϕ_o) Across Percentiles of the Occupation Task Distribution (1940 and 2000)



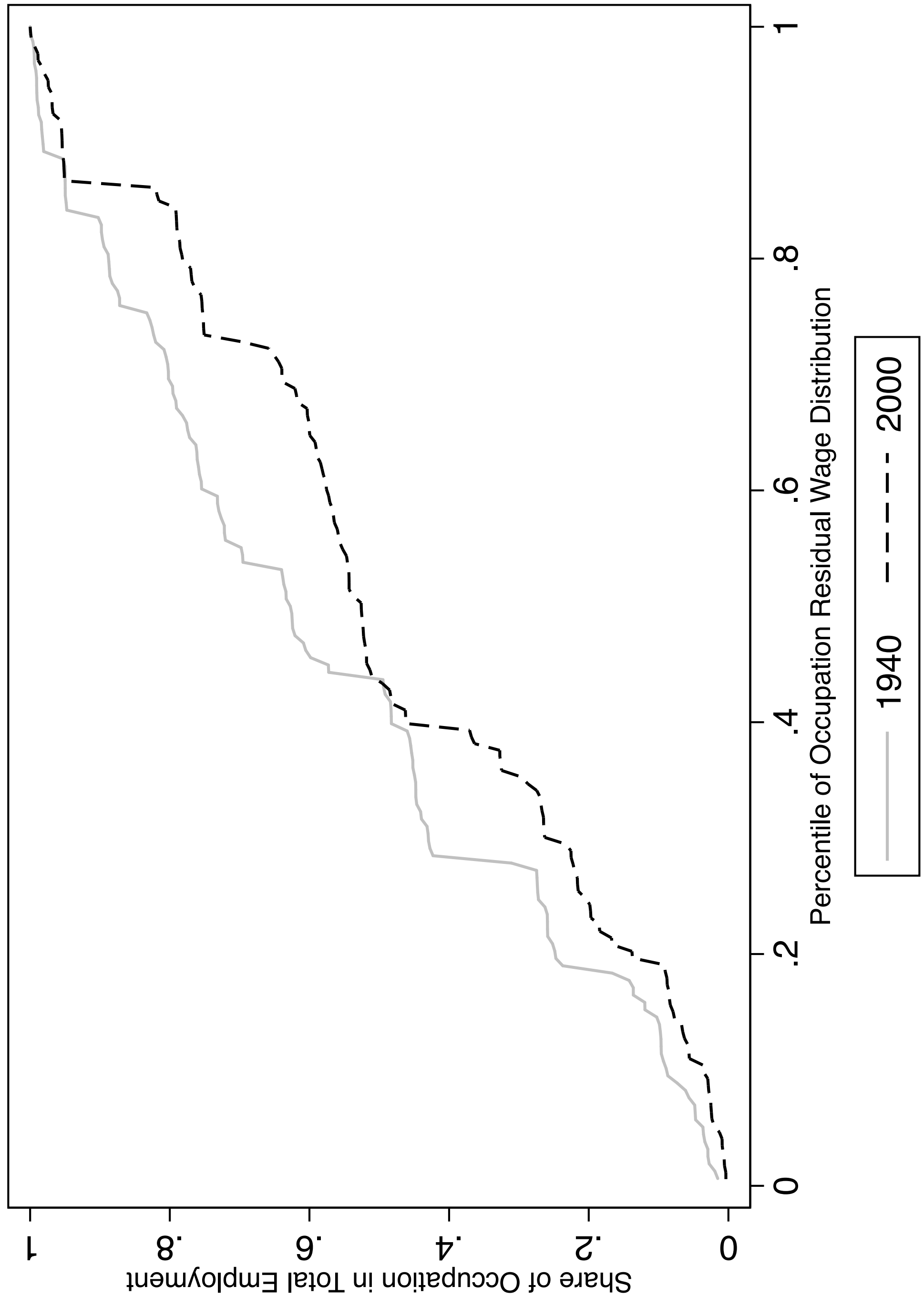
Note: On the horizontal axis, occupations are sorted according to their percentile of the occupation task distribution, as measured using the numerical scores from the Dictionary of Occupational Titles (DOTs) for 1991. "non-routine" is (non-routine analytic+non-routine interactive)/2; "routine" is (routine cognitive+routine manual)/2; and "manual" is non-routine manual. The vertical axis shows the cumulative distribution of human capital returns (ϕ_o from equation (A100) of this web appendix) for the sorted occupations (the cumulative sum of ϕ_o, scaled to add up to one).

Figure A4: Cumulative Distribution of Adjusted Task Effectiveness ($\mathbb{A}_o$) Across Percentiles of the Occupation Task Distribution (1940 and 2000)



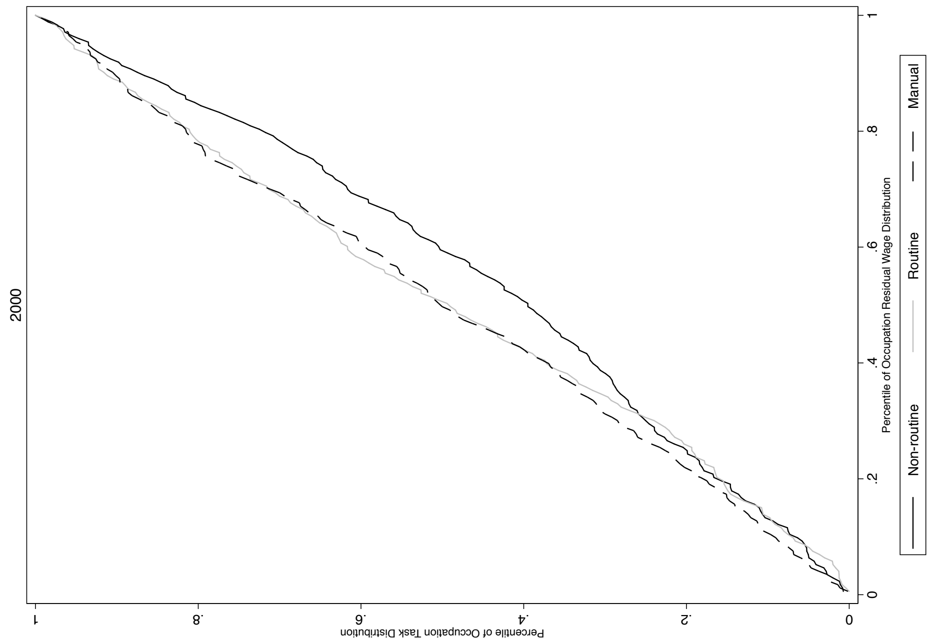
Note: On the horizontal axis, occupations are sorted according to their percentile of the occupation task distribution, as measured using the numerical scores from the Dictionary of Occupational Titles (DOTs) for 1991. "non-routine" is (non-routine analytic+non-routine interactive)/2; "routine" is (routine cognitive+routine manual)/2; and "manual" is non-routine manual. The vertical axis shows the cumulative distribution of occupation adjusted task effectiveness ($\mathbb{A}_o = \sum_{s=1}^{S} (E_{so}/E_o) \mathbb{A}_{so}$ from equation (A101) of this web appendix) for the sorted occupations (the cumulative sum of $\mathbb{A}_o$, scaled to add up to one).

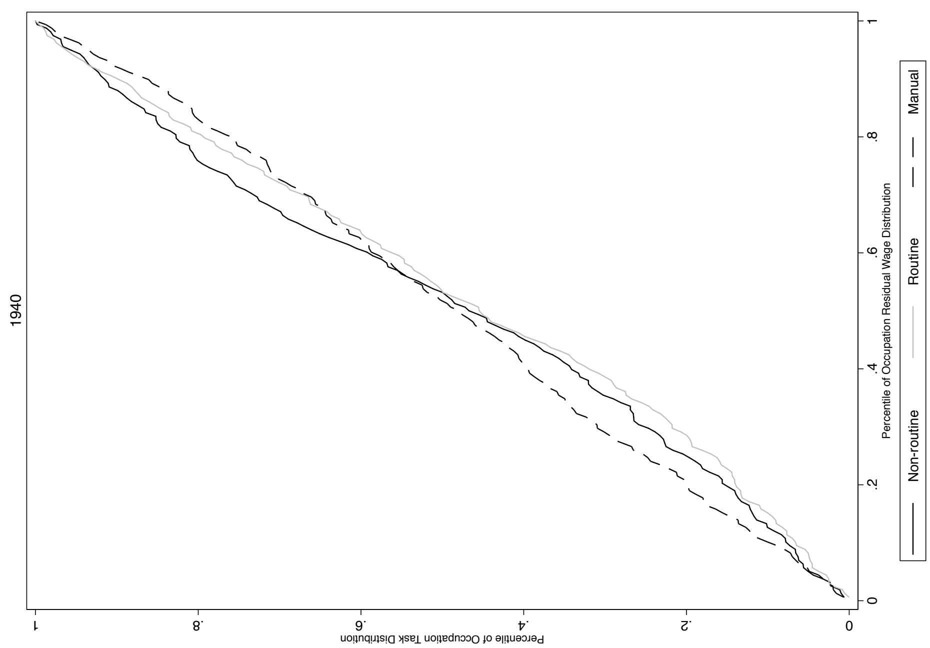
Figure A5: Cumulative Distribution of Occupation Employment Across Percentiles of the Occupation Residual Wage Distribution (1940 and 2000)



Note: On the horizontal axis, occupations are sorted in each year according to their percentile of the occupation residual wage distribution in that year. Residual wage inequality for each worker is measured as the estimated residual ($\hat{u}_{it}$) from the Mincer regression (equation (A102) of this web appendix) including controls for worker observables (education, age, gender and ethnicity). The residual wage for each occupation is measured as the average of the residual wage across workers within that occupation. The vertical axis shows the cumulative share of the sorted occupations in total employment in that year. Occupation employment and both worker wages and characteristics in the Mincer regression are measured using the IPUMs population census data.

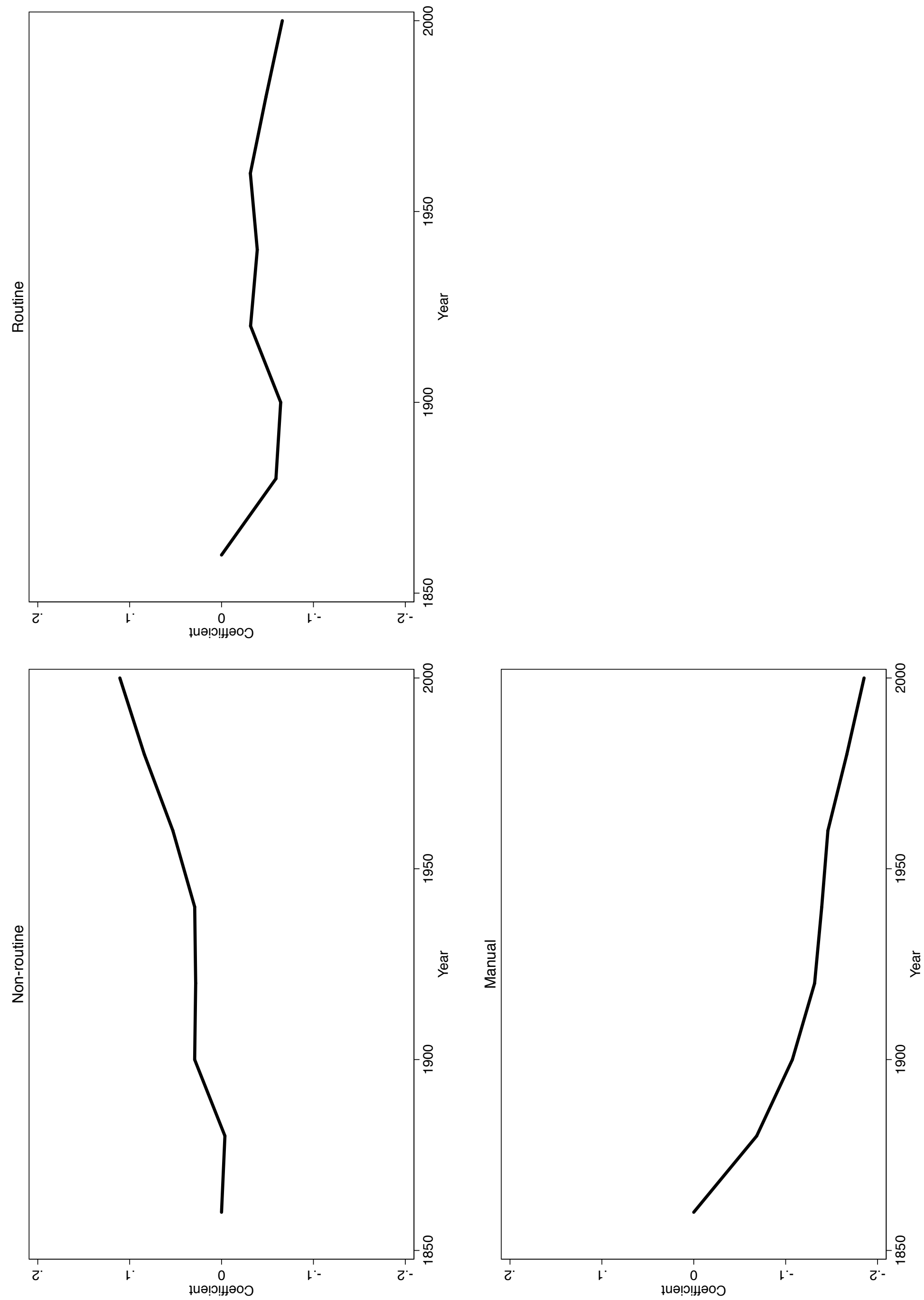
Figure A6: Cumulative Distribution of Tasks Across Percentiles of the Occupation Residual Wage Distribution (1940 and 2000)





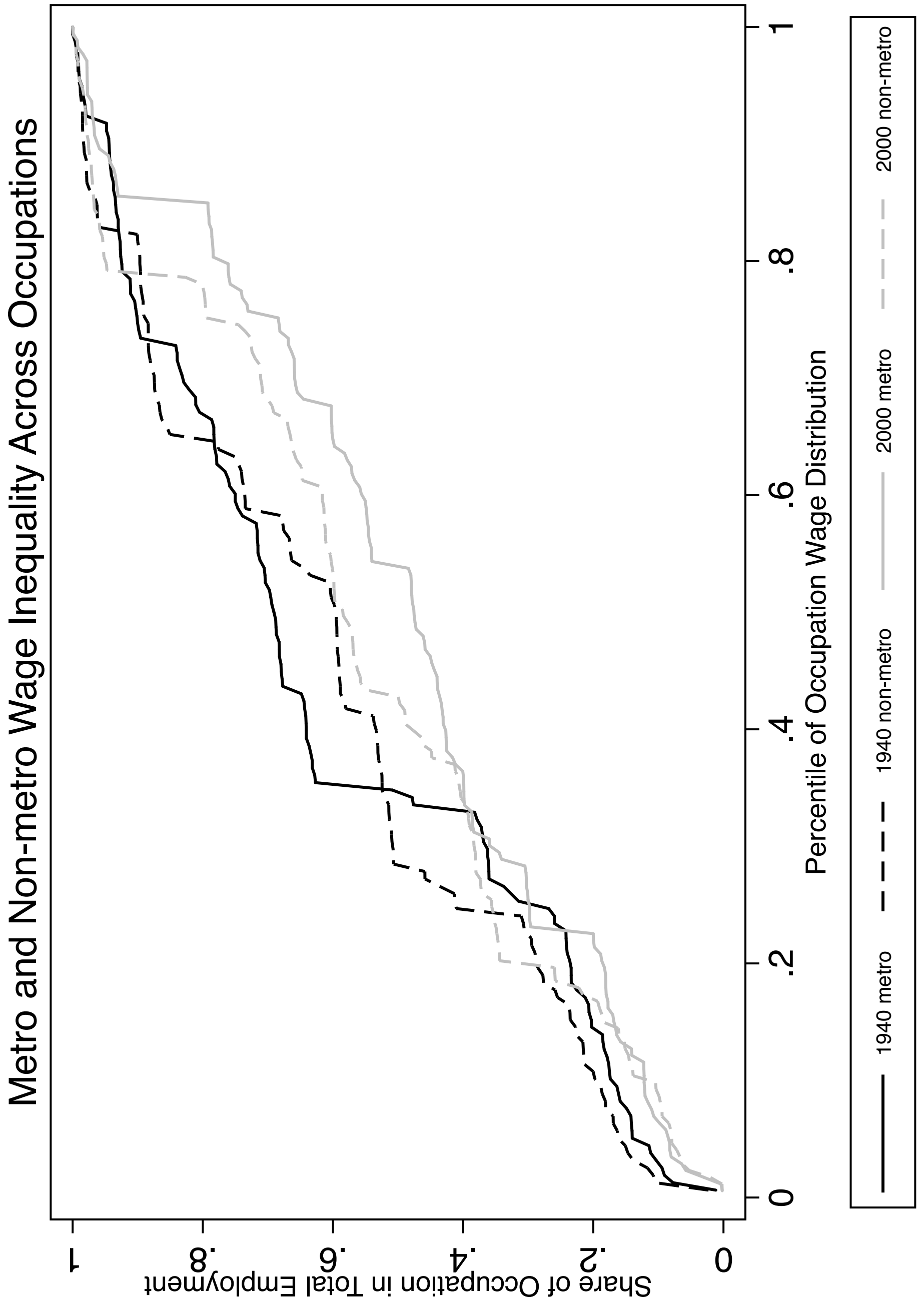
Note: On the horizontal axis, occupations are sorted in each year according to their percentile of the occupation residual wage distribution in that year. Residual wage inequality for each worker is measured as the estimated residual ($\hat{u}_{it}$) from the Mincer regression (A102) including controls for worker observables (education, age, gender and ethnicity). The residual wage for each occupation is measured as the average of the residual wage across workers within that occupation. The vertical axis shows the cumulative task distribution of the sorted occupations (the cumulative sum of the percentile numerical scores for the sorted occupations, scaled to add up to one). Numerical scores from the Dictionary of Occupational Titles (DOTs) for 1991. Each time-invariant numerical score is converted into the percentile of its distribution across occupations. “Non-routine” is (non-routine analytic+non-routine interactive)/2; “routine” is (routine cognitive+routine manual)/2; and “manual” is non-routine manual.

Figure A7: Task Input and Industry Office and Computing Machinery Use 1860-2000



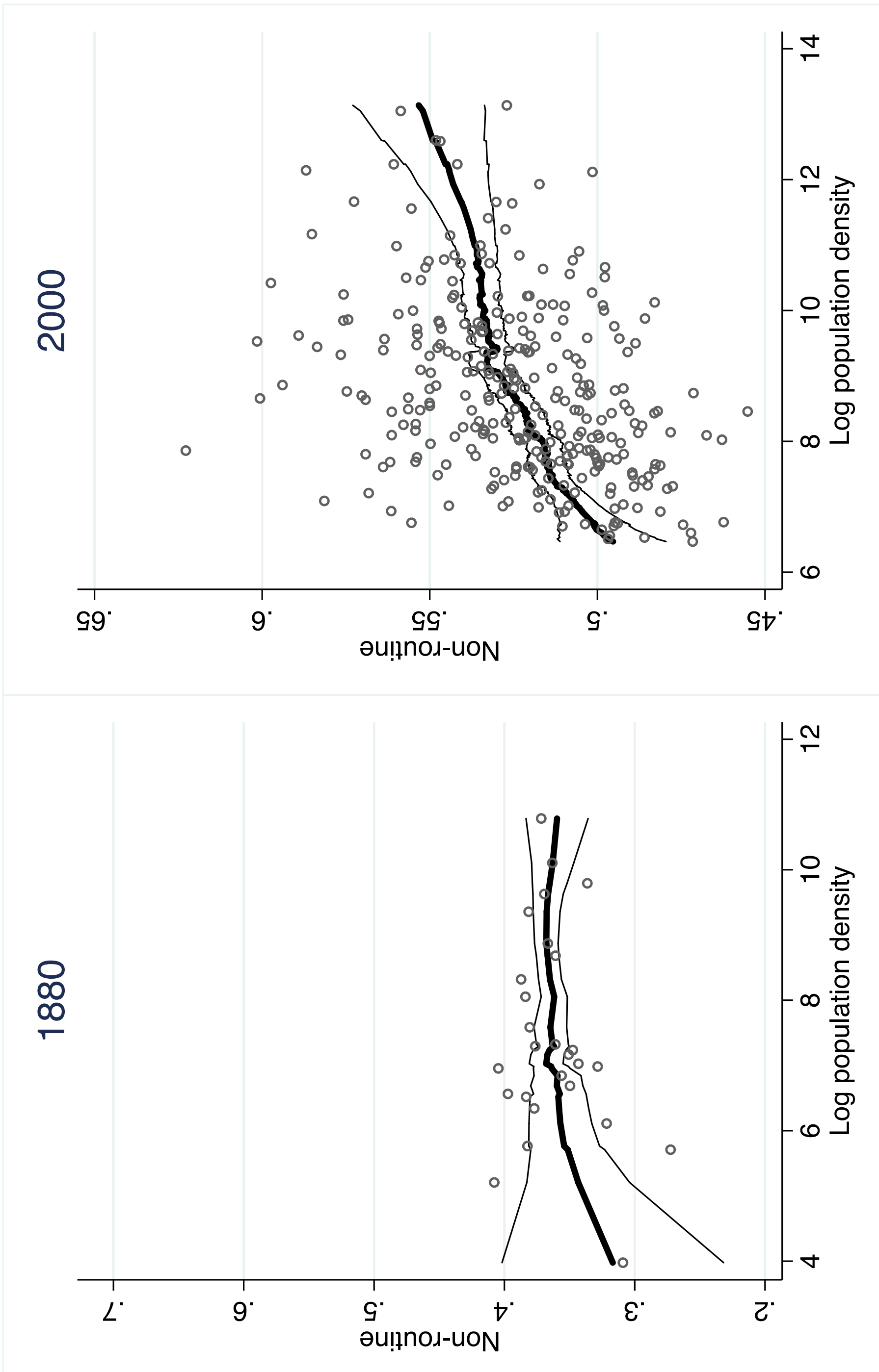
Note: Estimated coefficient (β_{tkm}) from the regression (A103) of industry inputs of task k (nonroutine, routine and manual) on a time-invariant measure of industry use of office and computing machinery m interacted with dummies for year t. Observations are industries and years. 1860 is the excluded year. Industry task inputs are the employment-weighted average of the non-routine, routine and manual numerical scores for each occupation from the Dictionary of Occupational Titles (DOTs) for 1991. Employment is measured using IPUMS population census data for each twenty-year interval from 1860-2000.

Figure A8: Cumulative Distribution of Occupation Employment Across Percentiles of the Occupation Residual Wage Distribution in Metro and Non-metro Areas (1940 and 2000)



Note: On the horizontal axis, occupations are sorted in each year according to their percentile of the occupation residual wage distribution for metro or non-metro areas in that year. Residual wage inequality for each worker is measured as the estimated residual ($\hat{u}_{it}$) from the Mincer regression (A102 of this web appendix) including controls for worker observables (education, age, gender and ethnicity). The residual wage for each occupation for metro areas is measured as the average of the residual wage across workers in metro areas within that occupation. The residual wage for each occupation for non-metro areas is measured analogously. The vertical axis shows the cumulative share of the sorted occupations in total employment in that year for metro and non-metro areas. Occupation employment, worker wages and characteristics in the Mincer regression and location by metro and non-metro area are measured using the IPUMs population census data.

Figure A9: Non-routine Task Inputs Across Metro Areas (1880 and 2000)



Note: Non-routine task input for each metro area is the employment-weighted sum of the non-routine numerical score for each occupation. Occupation numerical scores from the Dictionary of Occupational Titles (DOTs) for 1991. The time-invariant numerical score for each occupation is converted into the percentile of its distribution across occupations. Employment in each occupation and year in each metro area is measured using IPUMS population census data. Metro areas are defined based on time-varying boundaries from IPUMS. Thick solid line is the fitted value from a locally-weighted linear least squares regression. Thin solid lines are 95 percent point confidence intervals.

Table A1: Task Coefficients for the Probability of Being in a Metro Area

	1940	1960	1980	2000
Panel A: Education and Demographics				
Non-routine	-0.056	-0.006	0.008	0.014
	(0.076)	(0.036)	(0.032)	(0.008)
Routine	0.096^{**}	0.054^{*}	0.024	-0.006
	(0.047)	(0.033)	(0.031)	(0.008)
Manual	-0.128^{**}	-0.112^{***}	-0.102^{***}	-0.038^{***}
	(0.044)	(0.023)	(0.023)	(0.009)
Education & demographics	yes	yes	yes	yes
Panel B: Education, Demographics and Industry				
Non-routine	-0.050^{**}	0.004	0.021	0.014^{***}
	(0.025)	(0.015)	(0.014)	(0.005)
Routine	0.033	0.021	0.014	-0.006
	(0.030)	(0.016)	(0.015)	(0.004)
Manual	-0.074^{**}	-0.077^{***}	-0.071	-0.025^{***}
	(0.034)	(0.018)	(0.016)	(0.006)
Education & demographics	yes	yes	yes	yes
Industry Fixed Effects	yes	yes	yes	yes
Pabel C: Education, Demographics, Industry and Headquarters				
Non-routine	-0.087^{***}	-0.015	0.007	0.011^{*}
	(0.033)	(0.021)	(0.014)	(0.005)
Routine	0.048	0.032^{*}	0.024	-0.004
	(0.034)	(0.018)	(0.015)	(0.004)
Manual	-0.031	-0.047^{**}	-0.044^{***}	-0.021^{***}
	(0.032)	(0.021)	(0.016)	(0.006)
Education & demographics	yes	yes	yes	yes
Industry fixed effects	yes	yes	yes	yes
Headquarters occupations fixed effects	yes	yes	yes	yes

Note: Table reports the estimated coefficients (ς_t) on percentile task scores ($T_{o(i)}$) from equation (A104) for workers i in occupations o in year t. Each column in each panel corresponds to a separate regression for a separate year (column) and specification (panel). Panel A includes controls for education and demographics (education, age, gender and ethnicity). Panel B includes the controls for education and demographics and industry fixed effects. Panel C includes the controls for education and demographics, industry fixed effects and an indicator variable for headquarters occupations. See the data section of this appendix for further discussion of data sources and variable definitions. Standard errors in parentheses are heteroscedasticity robust and clustered on occupation. *** denotes significance at the 1 percent level; ** indicates significance at the 5 percent level; and * corresponds to significance at the 10 percent level.

Table A2: Top and Bottom Ten Verbs Most Concentrated in Metro Areas (1939 DOT)

Panel A: Verbs Most Strongly Correlated with Metro Area Employment Shares							
Rank	1880	1900	1920	1940	1960	1980	2000
1	Thread	Prevent	Detail	Keep	Interview	Interview	Advise
2	Slot	Flower	Interview	Accompany	Issue	Devise	Audit
3	Stitch	Need	Keep	Prevent	Accompany	Take	Question
4	Flower	Keep	Issue	Interview	Detail	Issue	Present
5	Help	Accompany	Address	Issue	Keep	Detail	Specialize
6	Observe	Detail	Devise	Address	Devise	Address	Report
7	Layer	Address	Prevent	Attend	Attend	Judge	Devise
8	Address	Thread	License	Detail	Address	Need	Promote
9	Straighten	Issue	Tell	Wage	Take	Tell	Formulate
10	Calk	Interview	Validate	Live	Judge	Validate	Implement

Panel B: Verbs Least Strongly Correlated with Metro Area Employment Shares							
Rank	1880	1900	1920	1940	1960	1980	2000
N-9	Sign	Promote	Promote	Proceed	Pack	Hole	Grease
N-8	Fund	Formulate	Advance	Allow	Skin	Splash	Wheel
N-7	Side	Lead	Fund	Formulate	Size	Employ	Jack
N-6	Build	Fund	Drum	Turn	Angle	Grease	Edge
N-5	Bore	Sign	Screw	Drum	Bundle	Lift	Employ
N-4	Program	Bore	Mark	Discharge	Shoot	Relieve	Lift
N-3	Work	Determine	Bore	Program	Work	Jack	Move
N-2	Drill	Program	Hang	Hang	Piece	Move	Remove
N-1	Part	Part	Weigh	Direct	Pile	Put	Advance
N	Mark	Mark	Formulate	Advance	Advance	Advance	Put

Note: Note: Panel A reports the ten verbs with the highest correlations with metro area employment shares, as measured by the ten verbs v with the most positive estimated coefficients in year t (α_{vt}) in equation (A105). Panel B reports the ten verbs with the most negative correlations with metro area employment shares, as measured by the ten verbs v with the smallest estimated coefficients in year t (α_{vt}) in equation (A105). Observations are sectors and occupations in each year. Metro area employment shares are measured as the share of employment within a sector-occupation that is located in a metro area in each year. The reported coefficients are standardized by variable standard deviations ("beta coefficients"). Verb frequencies for each occupation are measured from the occupational descriptions in the Dictionary of Occupational Titles (DOTs) for 1939 (VerbFreq_{vo} from equation (22) in the paper).

Table A3: Correlations With Independent Measures of Interactiveness

Unweighted Correlations		Weighted Correlations	
	Interactiveness		Interactiveness
Interactiveness	1	Interactiveness	1
Assisting and caring for others	0.23***	Assisting and caring for others	0.27***
Coaching and developing others	0.41***	Coaching and developing others	0.49***
Communicating with persons outside organization	0.51***	Communicating with persons outside organization	0.60***
Communicating with Supervisors, Peers, or Subordinates	0.46***	Communicating with Supervisors, Peers, or Subordinates	0.43***
Coordinating the work and activities of others	0.33***	Coordinating the work and activities of others	0.40***
Developing and building teams	0.41***	Developing and building teams	0.45***
Establishing and maintaining interpersonal relationships	0.63***	Establishing and maintaining interpersonal relationships	0.59***
Guiding, directing and motivating subordinates	0.35***	Guiding, directing and motivating subordinates	0.40***
Interpreting the meaning of information for others	0.51***	Interpreting the meaning of information for others	0.50***
Monitoring and controlling resources	0.35***	Monitoring and controlling resources	0.31***
Performing administrative activities	0.68***	Performing administrative activities	0.59***
Performing for or working directly with the public	0.29***	Performing for or working directly with the public	0.21***
Provide consultation and advice to others	0.55***	Provide consultation and advice to others	0.51***
Resolving conflict and negotiating with others	0.55***	Resolving conflict and negotiating with others	0.54***
Selling or influencing others	0.43***	Selling or influencing others	0.38***
Staffing organizational units	0.50***	Staffing organizational units	0.53***
Training and teaching others	0.39***	Training and teaching others	0.50***

Note: Table reports correlations across occupations between a measure of occupation interactiveness based on the frequency of use of tasks in Classes 4 and 5 of the thesaurus (the "Intellectual Faculties" and "Voluntary Powers" respectively) and independent measures of occupation interactiveness based on employee and employer surveys from O*NET. We consider all 17 subcategories of "Work Activities - Interacting with Others" from O*NET. Correlations reported across the sample of occupations in 2000. Weighted correlations are weighted by occupation employment in 2000. *** denotes significance at the 1 percent level; ** indicates significance at the 5 percent level; and * corresponds to significance at the 10 percent level.