# Web Appendix for Task Specialization in U.S. Cities from 1880-2000: Not for Publication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*[‡]

August 25, 2017

## 1 Introduction

This appendix contains the technical derivations of expressions and additional supplementary material for the main paper.

## 2 Theoretical Model

In this section, we outline a theoretical model that we use to interpret our empirical finding of an increased interactiveness of employment in urban areas relative to rural areas over time. The model explains the distribution of employment across occupations, sectors and locations. Despite allowing for a large number of locations and a rich geography of trade costs, the model remains tractable, because of the stochastic formulation of productivity differences across occupations, sectors and locations. The key predictions of the model are comparative statics with respect to the costs of trading the tasks produced by each occupation and the final goods produced by each sector. When these costs are large, all locations have similar employment structures across sectors, and all tasks within each sector are undertaken in the same location where the final good is produced. As the costs of trading final goods and tasks fall, locations specialize across sectors and across occupations within sectors according to their comparative advantage as determined by productivity differences. If densely-populated urban locations have a comparative advantage in interactive tasks relative to sparsely-populated rural locations, the model predicts that a fall in the costs of trading tasks leads to an increase in the interactiveness of employment within sectors in urban relative to rural areas.

[*] 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.

### 2.1 Preferences and Endowments

The economy consists of many locations indexed by $n \in N$. Each location $n$ is endowed with an exogenous supply of land $\bar{H}_n$. The economy as a whole is endowed with a measure of workers $\bar{L}$, who are perfectly mobile across locations.

Workers' preferences are defined over a goods consumption index ($C_n$) and residential land use ($H_n$) and are assumed to take the Cobb-Douglas form:[1]

$$U_n = \left(\frac{C_n}{\alpha}\right)^{\alpha}\left(\frac{H_n}{1-\alpha}\right)^{1-\alpha}, \qquad 0 < \alpha < 1. \tag{1}$$

The goods consumption index ($C_n$) is assumed to be a constant elasticity of substitution (CES) function of consumption indices for a number of sectors (e.g. Manufacturing, Services) indexed by $s \in S$:

$$C_n = \left[\sum_{s \in S} C_{ns}^{\frac{\beta-1}{\beta}}\right]^{\frac{\beta}{\beta-1}}, \tag{2}$$

where $\beta$ is the elasticity of substitution between sectors. Sectors can be either substitutes ($\beta > 1$) or complements in goods consumption ($0 < \beta < 1$), where the standard assumption in the literature on structural transformation in macroeconomics is complements (e.g. Ngai and Pissarides 2007, Yi and Zhang 2013).

The consumption index for each sector is in turn a CES function of consumption of a continuum of goods (e.g. Motor Vehicles, Drugs and Medicines) indexed by $j \in [0, 1]$:

$$C_{ns} = \left[\int_0^1 c_{ns}(j)^{\frac{\sigma_s-1}{\sigma_s}}\, dj\right]^{\frac{\sigma_s}{\sigma_s-1}}, \tag{3}$$

where the elasticity of substitution between goods $\sigma_s$ varies across sectors. While in the data we observe a finite number of goods within sectors, we adopt the theoretical assumption of a continuum of goods for reasons of tractability, because it enables us to make use of law of large numbers results in determining specialization at the sectoral level. Goods can be either substitutes ($\sigma_s > 1$) or complements ($0 < \sigma_s < 1$) and we can allow any ranking of the elasticities of substitution between goods and sectors, although the conventional assumption in such a nested CES structure is a higher elasticity of substitution at the more disaggregated level ($\sigma_s > \beta$).

Since the upper tier of utility is Cobb-Douglas, utility maximization implies that workers allocate constant shares of aggregate income to goods consumption and residential land use:

$$P_n C_n = \alpha v_n L_n, \tag{4}$$

$$r_n H_n = (1-\alpha)\, v_n L_n. \tag{5}$$

where $P_n$ is the price index dual to the goods consumption index ($C_n$); $v_n$ is income per worker; $L_n$ is the population of location $n$; and $r_n$ is the land rent.

Expenditure on residential land in each location is assumed to be redistributed lump-sum to residents of that location, as in Helpman (1998). Therefore aggregate income in each location ($v_n L_n$) equals payments to

[1] For empirical evidence using U.S. data in support of the constant expenditure share implied by the Cobb-Douglas functional form, see Davis and Ortalo-Magne (2011).

labor used in production ($w_n L_n$) plus expenditure on residential land ($r_n H_n = (1-\alpha) v_n L_n$):

$$v_n L_n = w_n L_n + (1-\alpha) v_n L_n = \frac{w_n L_n}{\alpha}, \tag{6}$$

where $w_n$ is the wage. Equilibrium land rents in each location ($r_n$) are determined by land market clearing, which requires that total land income equals total expenditure on land. Combining equilibrium expenditure on land (5), aggregate income (6) and land market clearing ($H_n = \bar{H}_n$), we obtain equilibrium land rents as a function of wages, population, the exogenous land supply and parameters:

$$r_n = \frac{1-\alpha}{\alpha} \frac{w_n L_n}{\overline{H}_n}. \tag{7}$$

## 2.2 Production

Goods are homogeneous in the sense that one unit of a given good is the same as any other unit of that good. Production occurs under conditions of perfect competition and constant returns to scale. The cost to a consumer in location $n$ of purchasing one unit of good $j$ within sector $s$ from location $i$ is therefore:

$$p_{nis}(j) = \frac{d_{nis} G_{is}(j)}{z_{is}(j)}, \tag{8}$$

where $d_{nis}$ are iceberg transport costs, such that $d_{nis} > 1$ must be shipped from location $i$ to location $n$ within sector $s$ in order for one unit to arrive; no arbitrage ensures that the triangle inequality $d_{nis} \leq d_{nks} d_{kis}$ is satisfied and we assume $d_{nns} = 1$; $z_{is}(j)$ is productivity for good $j$ within sector $s$ in location $i$; and $G_{is}(j)$ is the unit cost of the composite factor of production used for good $j$ within sector $s$ in location $i$, as determined below.

Final goods productivity is stochastic and modeled as in Eaton and Kortum (2002) and Costinot, Donaldson and Komunjer (2012). Final goods productivity for each good, sector and location is assumed to be drawn independently from a Fréchet distribution:[2]

$$F_{is}(z) = e^{-T_{is} L_{is}^{\eta_s} z^{\theta_s}}, \tag{9}$$

where the shape parameter $\theta_s > 1$ controls the dispersion of productivity across goods within each sector, which determines comparative advantage across goods. In contrast, the scale parameter ($T_{is} L_{is}^{\eta_s}$, where $\eta_s > 0$) determines average productivity within each sector for each location, which determines comparative advantage across sectors. We allow average productivity in a sector and location to be increasing in employment in that sector and location to capture agglomeration forces in the form of external economies of scale in final goods production (e.g. Ethier 1982).

We assume that the final good for each sector is produced using a number of stages of production, where each stage of production within a sector is supplied by a separate occupation indexed by $o \in O_s$ (e.g. Managers, Operatives). Output of good $j$ within sector $s$ in location $i$ ($Y_{is}(j)$) is a CES function of the inputs of each occupation ($X_{iso}(j)$):

$$Y_{is}(j) = \left[ \sum_{o \in O_s} X_{iso}(j)^{\frac{\mu_s - 1}{\mu_s}} \right]^{\frac{\mu_s}{\mu_s - 1}}, \tag{10}$$

[2] To simplify the exposition, we use $i$ to denote locations of production and $n$ to denote locations of consumption, except where otherwise indicated.

where $\mu_s$ is the elasticity of substitution between occupations and again we can allow occupations to be either substitutes ($\mu_s > 1$) or complements ($0 < \mu_s < 1$). Under our assumption of a CES technology, the value marginal product of each occupation becomes infinite as the input of that occupation converges towards zero. Therefore the inputs of all occupations $o \in O_s$ within each sector are used in positive amounts. But sectors can differ in their set of occupations $O_s$ and firms within each sector can adjust the proportions in which the inputs of these occupations are used depending their cost.

Workers within each occupation perform a continuum of tasks $t \in [0, 1]$ as in Grossman and Rossi-Hansberg (2008) (e.g. as captured by the verbs Advising, Typing, Stretching, Stamping in our empirical analysis). The input for occupation $o$ and good $j$ within sector $s$ and location $i$ ($X_{iso}(j)$) is a CES function of the inputs for these tasks ($x_{iso}(j, t)$):

$$X_{iso}(j) = \left[\int_0^1 x_{iso}(j,t)^{\frac{\nu_{so}-1}{\nu_{so}}} dt\right]^{\frac{\nu_{so}}{\nu_{so}-1}} \tag{11}$$

where the elasticity of substitution between tasks $\nu_{so}$ varies across sectors and occupations. While in the data we observe a finite number of tasks within occupations, we adopt the theoretical assumption of a continuum of tasks for reasons of tractability, because it enables us to make use of law of large numbers results in determining specialization at the occupational level.[3] We allow tasks within occupations to be either substitutes ($\nu_{so} > 1$) or complements ($0 < \nu_{so} < 1$), and we can consider any ranking of the elasticities of substitution between tasks and occupations, although the conventional assumption in such a nested CES structure is again a higher elasticity of substitution at the more disaggregated level ($\nu_{so} > \mu_s$). Under our assumption of a CES technology, the value marginal product of each task also becomes infinite as the use of that task converges towards zero. Therefore all tasks within each occupation are used in positive amounts, although firms can adjust the proportions in which these tasks are used depending on their cost.[4]

Tasks are performed by labor using a constant returns to scale technology and can be traded between locations. For example, product design can be undertaken in one location, while production and assembly occur in another location. The cost to a firm in location $n$ of sourcing a task $t$ from location $i$ within occupation $o$ and sector $s$ is:

$$g_{niso}(j,t) = \frac{\tau_{niso} w_i}{a_{iso}(j,t)}, \tag{12}$$

where $w_i$ is the wage; $\tau_{niso}$ are iceberg communication costs, such that $\tau_{niso} > 1$ units of the task must be performed in location $i$ in order for one unit to be completed in location $n$ for occupation $o$ and sector $s$; no arbitrage ensures that the triangle inequality $\tau_{niso} \leq \tau_{nkso}\tau_{kiso}$ is satisfied and we assume $\tau_{nnso} = 1$; $a_{iso}(j, t)$ is productivity for task $t$ and good $j$ within occupation $o$ and sector $s$ in location $i$.

Input productivity for each task, occupation, sector and location is also stochastic and is assumed to be drawn independently from a Fréchet distribution:

$$\mathcal{F}_{iso} = e^{-U_{iso} L_{iso}^{\chi_{so}} a^{-\epsilon_{so}}}, \tag{13}$$

[3] To reduce the notational burden, we assume the same $[0, 1]$ interval of tasks for all occupations, but it is straightforward to allow this interval to vary across occupations.

[4] While we interpret production as being undertaken by workers in occupations that perform many tasks, an equivalent interpretation is that each occupation corresponds to a stage of production and each task corresponds to an intermediate input within that stage of production.

where the shape parameter $\epsilon_{so} > 1$ controls the dispersion of productivity across tasks within occupations, which determines comparative advantage across tasks. In contrast, the scale parameter ($U_{iso}L_{iso}^{\chi_{so}} > 0$, where $\chi_{so} > 0$) controls average productivity within each occupation, which determines comparative advantage across occupations. We allow average productivity in an occupation, sector and location to be increasing in employment in that occupation, sector and location ($\chi_{so} > 0$) to capture external economies of scale in task production (e.g. Grossman and Rossi-Hansberg 2012).

## 2.3 Trade in Tasks and Input Costs

### 2.3.1 Locations' Shares of Costs within an Occupation and Sector

Firms within a given location $n$ source each task $t$ within an occupation $o$, good $j$ and sector $s$ from the lowest cost source of supply for that task:

$$g_{nso}(j,t) = \min\left\{g_{niso}(j,t); i \in N\right\}.$$

Using task prices (12) and the Fréchet distribution of input productivities (13), the distribution of task prices in country $n$ for goods sourced from country $i$ in occupation $o$ within sector $s$ is:

$$\mathcal{F}_{niso}(g) = \Pr[g_{niso} \leq g] = 1 - \mathcal{F}_{iso}\left(\frac{w_i \tau_{niso}}{g}\right),$$

$$\mathcal{F}_{niso}(g) = 1 - e^{-U_{iso}L_{iso}^{\chi_{so}}(\tau_{niso}w_i)^{-\varepsilon_{so}}g^{\varepsilon_{so}}}. \tag{14}$$

Tasks are sourced from the lowest-cost supplier and the distribution of minimum task prices in country $n$ in occupation $o$ within sector $s$ is:

$$\mathcal{F}_{nso}(g) = 1 - \prod_{i\in N}\left[1 - \mathcal{F}_{niso}(g)\right] = 1 - e^{-\Psi_{nso}g^{\varepsilon_{so}}}, \tag{15}$$

$$\Psi_{nso} \equiv \sum_{i\in N} U_{iso}L_{iso}^{\chi_{so}}\left(\tau_{niso}w_i\right)^{-\varepsilon_{so}}, \tag{16}$$

Since tasks are sourced from the lowest-cost supplier, the probability that location $n$ sources a task $t$ within occupation $o$ and sector $s$ from location $i$ is:

$$\begin{aligned}\lambda_{niso} &= \Pr\left[g_{niso}(t) \leq \min\left\{g_{nkso}(t)\right\}; k \neq i\right] \\ &= \int_0^\infty \prod_{k\neq i}\left[1 - \mathcal{F}_{nkso}(g)\right] d\mathcal{F}_{niso}(g).\end{aligned}$$

Using the bilateral price distribution (14), the probability that location $n$ sources a task $t$ from location $i$ within occupation $o$ and sector $s$ is:

$$\lambda_{niso} = \frac{U_{iso}L_{iso}^{\chi_{so}}\left(\tau_{niso}w_i\right)^{-\epsilon_{so}}}{\sum_{k\in N} U_{kso}L_{kso}^{\chi_{so}}\left(\tau_{nkso}w_k\right)^{-\epsilon_{so}}}. \tag{17}$$

Since the Fréchet distribution is unbounded from above, each location draws an arbitrarily high input productivity for a positive measure of tasks. To allow for the possibility that a location may not have positive employment in an occupation $o$ and sector $s$, we take $\lim U_{iso} \to 0$, in which case the location's employment in that occupation and sector converges to zero. Similarly, to allow for the possibility that an occupation $o$ may not be traded, we take $\lim d_{nis} \to \infty$, in which case trade in that occupation converges to zero.

Another implication of the Fréchet distribution of input productivities is that the distribution of task prices in location $n$ for tasks actually sourced from another location $i$ is independent of the identity of the location $i$ and equal to the distribution of minimum prices in location $n$. To derive this result, note that the distribution of task prices in location $n$ conditional on sourcing tasks from location $i$ is:

$$\frac{1}{\lambda_{niso}} \int_0^g \prod_{k \neq i} \left[1 - \mathcal{F}_{nkso}\left(g'\right)\right] d\mathcal{F}_{niso}\left(g'\right) = 1 - e^{-\Psi_{nso} g^{\varepsilon_{so}}} = \mathcal{F}_{nso}\left(g\right),$$

where we have used the bilateral and multilateral price distributions, (14) and (15) respectively. Intuitively, under the assumption of a Fréchet distribution of input productivity, a source location $i$ with a higher scale parameter ($U_{iso} L_{iso}^{\chi_{so}}$), and hence a higher average input productivity, expands on the extensive margin of the number of tasks supplied exactly to the point at which the distribution of prices for the tasks it actually sells in market $n$ is the same as destination $n$'s distribution of minimum prices.

Since the distribution of prices in location $n$ for goods actually purchased is the same across all source locations $i$, it follows that the share of location $n$'s expenditure on products sourced from another location $i$ within occupation $o$ and sector $s$ is equal to the probability of sourcing a task from that location ($\lambda_{niso}$). Therefore the share of location $n$'s expenditure on tasks sourced from another location $i$ within occupation $o$ and sector $s$ is given by (17).

### 2.3.2 Occupations' Shares of Costs

We begin by determining the cost function for occupation $o$ and sector $s$ in location $n$ using the distribution of minimum task prices (15):

$$\begin{aligned} G_{nso} &= \left[\int_0^1 g_{nso}\left(t\right)^{1-\nu_{so}} dt\right]^{\frac{1}{1-\nu_{so}}}, \\ &= \left[\int_0^\infty g_{nso}^{1-\nu_{so}} d\mathcal{F}_{nso}\left(g\right)\right]^{\frac{1}{1-\nu_{so}}}, \\ &= \left[\int_0^\infty \varepsilon_{so} \Psi_{nso} g^{\varepsilon_{so}-\nu_{so}} e^{-\Psi_{nso} g^{\varepsilon_{so}}} dg\right]^{\frac{1}{1-\nu_{so}}}. \end{aligned}$$

Using the following change of variable:

$$\tilde{g} = \Psi_{nso} g^{\varepsilon_{so}},$$

$$\Rightarrow \qquad g = \left(\frac{\tilde{g}}{\Psi_{nso}}\right)^{\frac{1}{\varepsilon_{so}}}, \qquad dg = \frac{1}{\theta_K}\left(\frac{\tilde{g}}{\Psi_{nso}}\right)^{\frac{1-\varepsilon_{so}}{\varepsilon_{so}}} \frac{1}{\Psi_{nso}} d\tilde{g},$$

we obtain:

$$G_{nso} = \Psi_{nso}^{-1/\varepsilon_{so}} \left[\int_0^\infty \tilde{g}^{(1-\nu_{so})/\varepsilon_{so}} e^{-\tilde{g}} d\tilde{g}\right]^{\frac{1}{1-\nu_{so}}},$$

which yields the following expression for the cost function for occupation $o$ and sector $s$ in location $n$:

$$G_{nso} = \gamma_{so} \Psi_{nso}^{-1/\varepsilon_{so}} = \gamma_{so} \left[\sum_{i \in N} U_{iso} L_{iso}^{\chi_{so}} \left(\tau_{niso} w_i\right)^{-\varepsilon_{so}}\right]^{-1/\varepsilon_{so}}, \tag{18}$$

$$\text{where} \qquad \gamma_{so} \equiv \left[\Gamma\left(\frac{\varepsilon_{so}+1-\nu_{so}}{\varepsilon_{so}}\right)\right]^{\frac{1}{1-\nu_{so}}},$$

where $\Gamma(\cdot)$ is the gamma function.

Together the cost share (17) and cost function (18) imply that the unit cost for occupation $o$ and sector $s$ in location $n$ also can be written as:

$$G_{nso} = \gamma_{so}\left(\frac{U_{nso}L_{nso}^{\chi_{so}}}{\lambda_{nnso}}\right)^{-\frac{1}{\epsilon_{so}}} w_n. \tag{19}$$

Given these unit costs for each occupation, we now solve for the overall unit cost for sector $s$ in location $n$. From the CES production technology (10), the overall unit cost is:

$$G_{ns} = \left[\sum_{o\in O_s} G_{nso}^{1-\mu_s}\right]^{\frac{1}{1-\mu_s}},$$

which using the unit costs for each occupation (19) can be written as:

$$G_{ns} = \left[\sum_{o\in O_s} \gamma_{so}^{1-\mu_s}\left(\frac{U_{nso}L_{nso}^{\chi_{so}}}{\lambda_{nnso}}\right)^{-\frac{1-\mu_s}{\epsilon_{so}}}\right]^{\frac{1}{1-\mu_s}} w_n. \tag{20}$$

The CES production technology (10) also implies that the share of occupation $o$ in unit costs within sector $s$ in location $n$ is:

$$e_{nso} = \frac{G_{nso}^{1-\mu_s}}{\sum_{m\in O_s} G_{nsm}^{1-\mu_s}},$$

which using the unit costs for each occupation (19) can be written as the expression in the paper:

$$e_{nso} = \frac{\gamma_{so}^{1-\mu_s}\left(\frac{U_{nso}L_{nso}^{\chi_{so}}}{\lambda_{nnso}}\right)^{-\frac{1-\mu_s}{\epsilon_{so}}}}{\sum_{m\in O_s}\gamma_{sm}^{1-\mu_s}\left(\frac{U_{nsm}L_{nsm}^{\chi_{so}}}{\lambda_{nnsm}}\right)^{-\frac{1-\mu_s}{\epsilon_{sm}}}}. \tag{21}$$

## 2.4 Trade in Final Goods and Price Indices

### 2.4.1 Locations' Shares of Sectoral Expenditure

Consumers within a given location $n$ source each final good $j$ within a sector $s$ from the lowest cost source of supply for that final good:

$$p_{ns}(j) = \min\left\{p_{nis}(j); i \in N\right\}.$$

Using final goods prices (8) and the Fréchet distribution of final goods productivities (9), the distribution of goods prices in country $n$ for goods sourced from country $i$ within sector $s$ is:

$$F_{nis}(p) = \Pr[p_{nis} \le p] = 1 - F_{is}\left(\frac{d_{nis}G_{is}}{p}\right),$$

$$F_{nis}(p) = 1 - e^{-T_{is}L_{is}^{\eta_s}(d_{nis}G_{is})^{-\theta_s}p^{\theta_s}}. \tag{22}$$

Final goods are sourced from the lowest-price supplier and the distribution of minimum final goods prices in country $n$ in within sector $s$ is:

$$F_{ns}(p) = 1 - \prod_{i \in N} \left[1 - F_{nis}(p)\right] = 1 - e^{-\Lambda_{ns} p^{\theta_s}}, \tag{23}$$

$$\Lambda_{ns} \equiv \sum_{i \in N} T_{is} L_{is}^{\eta_s} \left(d_{nis} G_{is}\right)^{-\theta_s}. \tag{24}$$

Since final goods are sourced from the lowest-price supplier, the probability that location $n$ sources a final good $j$ within sector $s$ from location $i$ is:

$$\begin{aligned} \pi_{nis} &= \Pr\left[p_{nis}(j) \leq \min\left\{p_{nks}(j)\right\}; k \neq i\right], \\ &= \int_0^\infty \prod_{k \neq i} \left[1 - F_{nks}(p)\right] dF_{nis}(p). \end{aligned}$$

Using the bilateral price distribution (22), the probability that location $n$ sources a final good $j$ from location $i$ within sector $s$ is:

$$\pi_{nis} = \frac{T_{is} L_{is}^{\eta_s} \left(d_{nis} G_{is}\right)^{-\theta_s}}{\sum_{k \in N} T_{ks} L_{ks}^{\eta_s} \left(d_{nks} G_{is}\right)^{-\theta_s}},$$

which can be in turn re-written as the following expression in the paper:

$$\pi_{nis} = \frac{T_{is} L_{is}^{\eta_s} \left(d_{nis} \Phi_{is} w_i\right)^{-\theta_s}}{\sum_{k \in N} T_{ks} L_{ks}^{\eta_s} \left(d_{nks} \Phi_{is} w_i\right)^{-\theta_s}}, \tag{25}$$

where from the previous subsection:

$$\Phi_{is} = \left[\sum_{o \in O_s} \gamma_{so}^{1-\mu_s} \left(\frac{U_{iso} L_{iso}^{\chi_{so}}}{\lambda_{iiso}}\right)^{-\frac{1-\mu_s}{\epsilon_{so}}}\right]^{\frac{1}{1-\mu_s}}.$$

Since the Fréchet distribution is unbounded from above, each location draws an arbitrarily high final goods productivity for a positive measure of final goods. To allow for the possibility that a location may not have positive employment in a sector $s$, we take $\lim T_{is} \to 0$, in which case the location's employment in that sector converges to zero. Similarly, to allow for the possibility that a sector $s$ may not be traded, we take $\lim d_{nis} \to \infty$, in which case trade in that sector converges to zero.

Another implication of the Fréchet distribution of final goods productivities is that the distribution of final goods prices in location $n$ for goods actually sourced from another location $i$ is independent of the identity of the location $i$ and equal to the distribution of minimum prices in location $n$. To derive this result, note that the distribution of final goods prices in location $n$ conditional on sourcing goods from location $i$ is:

$$\frac{1}{\pi_{nis}} \int_0^p \prod_{k \neq i} \left[1 - F_{nks}(p')\right] dF_{nis}(p') = 1 - e^{-\Lambda_{ns} p^{\theta_s}} = F_{ns}(p),$$

where we have used the bilateral and multilateral price distributions, (22) and (23) respectively. Intuitively, under the assumption of a Fréchet distribution of final goods productivity, a source location $i$ with a higher scale parameter ($T_{is} L_{is}^{\eta_s}$), and hence a higher average final goods productivity, expands on the extensive margin of the number of final goods supplied exactly to the point at which the distribution of prices for the goods it actually sells in market $n$ is the same as destination $n$'s distribution of minimum prices.

Since the distribution of prices in location $n$ for goods actually purchased is the same across all source locations $i$, it follows that the share of location $n$'s expenditure on final goods sourced from another location $i$ within sector $s$ is equal to the probability of sourcing a final good from that location ($\pi_{nis}$). Therefore the share of location $n$'s expenditure on final goods sourced from another location $i$ within sector $s$ is given by (25).

### 2.4.2 Sectors' Shares of Expenditure

We begin by determining the price index for sector $s$ in location $n$ using the distribution of minimum final goods prices (23):

$$\begin{aligned} P_{ns} &= \left[\int_0^1 p_{ns}(j)^{1-\sigma_s}\, dj\right]^{\frac{1}{1-\sigma_s}}, \\ &= \left[\int_0^\infty p_{ns}^{1-\sigma_s} dF_{ns}(p)\right]^{\frac{1}{1-\sigma_s}}, \\ &= \left[\int_0^\infty \theta_s \Lambda_{ns} p^{\theta_s-\sigma_s} e^{-\Lambda_{ns}p^{\theta_s}} dp\right]^{\frac{1}{1-\sigma_s}}. \end{aligned}$$

Using the following change of variable:

$$\tilde{p} = \Lambda_{ns} g^{\theta_s},$$

$$\Rightarrow \qquad p = \left(\frac{\tilde{p}}{\Lambda_{ns}}\right)^{\frac{1}{\theta_s}}, \qquad dp = \frac{1}{\theta_s}\left(\frac{\tilde{p}}{\Lambda_{ns}}\right)^{\frac{1-\theta_s}{\theta_s}} \frac{1}{\Lambda_{ns}} d\tilde{p},$$

we obtain:

$$P_{ns} = \Lambda_{ns}^{-1/\theta_s}\left[\int_0^\infty \tilde{p}^{(1-\sigma_s)/\theta_s} e^{-\tilde{p}} d\tilde{p}\right]^{\frac{1}{1-\sigma_s}},$$

which yields the following expression for the price index for sector $s$ in location $n$:

$$P_{ns} = \kappa_s \Lambda_{ns}^{-1/\theta_s} = \kappa_s \left[\sum_{i\in N} T_{is} L_{is}^{\eta_s} \left(d_{nis} G_{is}\right)^{-\theta_s}\right]^{-1/\theta_s},$$

$$\text{where} \qquad \kappa_s \equiv \left[\Gamma\left(\frac{\theta_s+1-\sigma_s}{\theta_s}\right)\right]^{\frac{1}{1-\sigma_s}},$$

where $\Gamma(\cdot)$ is the gamma function. This expression for the sectoral price index can be in turn re-written as:

$$P_{ns} = \kappa_s \left[\sum_{k\in N} T_{ks} L_{ks}^{\eta_s} \left(d_{nks} \Phi_{ks} w_k\right)^{-\theta_s}\right]^{-\frac{1}{\theta_s}}, \tag{26}$$

where $\Phi_{ks}$ is defined above.

Together the expenditure share (25) and cost function (26) imply that the price index for sector $s$ in location $n$ also can be written as:

$$P_{ns} = \kappa_s \left(\frac{T_{ns} L_{ns}^{\eta_s}}{\pi_{nns}}\right)^{-\frac{1}{\theta_s}} \Phi_{ns} w_n. \tag{27}$$

Given this price index for each sector, we now solve for the overall goods consumption price index in location $n$. From the CES goods consumption index (2), the corresponding dual price index is:

$$P_n = \left[\sum_{s\in S} P_{ns}^{1-\beta}\right]^{\frac{1}{1-\beta}},$$

which using the price index for each sector (27) can be written as:

$$P_n = \left[\sum_{s\in S} \kappa_s^{1-\beta} \left(\frac{T_{ns}L_{ns}^{\eta_s}}{\pi_{nns}}\right)^{-\frac{1-\beta}{\theta_s}} \Phi_{ns}^{1-\beta}\right]^{\frac{1}{1-\beta}} w_n, \tag{28}$$

The CES goods consumption index (2) also implies that the share of sector $s$ in aggregate goods consumption expenditure is:

$$E_{ns} = \frac{P_{ns}^{1-\beta}}{\sum_{r\in S} P_{nr}^{1-\beta}},$$

which using the price index for each sector (27) can be written as the expression in the paper:

$$E_{ns} = \frac{\kappa_s^{1-\beta} \left(\frac{T_{ns}L_{ns}^{\eta_s}}{\pi_{nns}}\right)^{-\frac{1-\beta}{\theta_s}} \Phi_{ns}^{1-\beta}}{\sum_{r\in S} \kappa_r^{1-\beta} \left(\frac{T_{nr}L_{nr}^{\eta_r}}{\pi_{nnr}}\right)^{-\frac{1-\beta}{\theta_r}} \Phi_{nr}^{1-\beta}}. \tag{29}$$

## 2.5 Population Mobility

Population mobility implies that workers must receive the same indirect utility in all populated locations:

$$V_n = \frac{v_n}{P_n^{\alpha} r_n^{1-\alpha}} = \bar{V}. \tag{30}$$

Using land market clearing (7), this population mobility condition becomes:

$$V_n = \frac{v_n}{P_n^{\alpha}\left(\frac{1-\alpha}{\alpha}\frac{w_n L_n}{\bar{H}_n}\right)^{1-\alpha}} = \bar{V},$$

which using the equality of income and expenditure (6) becomes:

$$V_n = \frac{w_n^{\alpha}}{\alpha P_n^{\alpha}\left(\frac{1-\alpha}{\alpha}\frac{L_n}{\bar{H}_n}\right)^{1-\alpha}} = \bar{V},$$

which using the aggregate price index (28) can be written as:

$$V_n = \frac{\left(\sum_{s\in S} \kappa_s^{-(1-\beta)} \left(\frac{T_{ns}L_{ns}^{\eta_s}}{\pi_{nns}}\right)^{\frac{1-\beta}{\theta_s}} \Phi_{ns}^{-(1-\beta)}\right)^{\frac{\alpha}{1-\beta}} \bar{H}_n^{1-\alpha}}{\alpha\left(\frac{1-\alpha}{\alpha}\right)^{1-\alpha} L_n^{(1-\alpha)}} = \bar{V}. \tag{31}$$

Re-arranging the above population mobility condition, we obtain the following expression for equilibrium population:

$$L_n = \frac{\left[\sum_{s\in S} \kappa_s^{-(1-\beta)} \left(\frac{T_{ns}L_{ns}^{\eta_s}}{\pi_{nns}}\right)^{\frac{1-\beta}{\theta_s}} \left[\sum_{o\in O_s} \gamma_{so}^{-(1-\mu_s)} \left(\frac{U_{nso}L_{nso}^{\chi_{so}}}{\lambda_{nnso}}\right)^{\frac{1-\mu_s}{\epsilon_{so}}}\right]^{\frac{1-\beta}{1-\mu_s}}\right]^{\frac{\alpha}{(1-\alpha)(1-\beta)}} \bar{H}_n}{\alpha^{\frac{1}{1-\alpha}}\left(\frac{1-\alpha}{\alpha}\right)\bar{V}^{\frac{1}{1-\alpha}}}, \tag{32}$$

where labor market clearing requires:

$$\sum_{n \in N} L_n = \bar{L}. \tag{33}$$

## 2.6 Welfare Gains from Trade

Rearranging the population mobility condition (31), indirect utility can be written as:

$$V_n = \frac{\left( \sum_{s \in S} \kappa_s^{-(1-\beta)} \left( \frac{T_{ns} L_{ns}^{\eta_s}}{\pi_{nns}} \right)^{\frac{1-\beta}{\theta_s}} \left[ \sum_{o \in O_s} \gamma_{so}^{-(1-\mu_s)} \left( \frac{U_{nso} L_{nso}^{\chi_{so}}}{\lambda_{nnso}} \right)^{\frac{1-\mu_s}{\epsilon_{so}}} \right]^{\frac{1-\beta}{1-\mu_s}} \right)^{\frac{\alpha}{1-\beta}} \bar{H}_n^{1-\alpha}}{\alpha \left( \frac{1-\alpha}{\alpha} \right)^{1-\alpha} L_n^{(1-\alpha)}}.$$

where the special case of no trade in final goods or tasks ($\lim_{d_{nis} \to \infty}$ and $\lim_{\tau_{niso} \to \infty}$) implies $\pi_{nns} = \lambda_{nnso} = 1$ for all $s, o$.

Therefore the welfare gains from trade depend on three components in this model. First, there are welfare gains from trade in final goods ($0 < \pi_{nns} < 1$). Second, there are welfare gains from trade in tasks ($0 < \lambda_{nnso} < 1$). Third, population mobility equalizes indirect utility across locations in both the closed and open economy. Therefore, if the opening of trade has uneven effects on the welfare of locations, population adjusts to ensure real wage equalization. It follows that the welfare gains from trade are the same for all locations and also depend on endogenous population ($L_n$). To the extent that trade in tasks ($0 < \lambda_{nnso} < 1$) is not fully captured in standard data on trade in goods, the model implies that measures of the welfare gains from trade based on these standard data will understate the true magnitude of the welfare gains from trade.

## 2.7 Wages and Employment

Wages in each location can be determined from the equality between a location's labor income and expenditure on tasks performed in that location:

$$w_i L_i = \sum_s \sum_o \sum_n \sum_k \lambda_{kiso} e_{kso} \left[ \pi_{nks} E_{ns} w_n L_n \right], \tag{34}$$

where the term inside the square parentheses on the right-hand side is market $n$'s population ($L_n$) times its wage ($w_n$) times the fraction of income that is allocated to final goods supplied by location $k$ in sector $s$ ($\pi_{nks} E_{ns}$); the term outside the square parentheses is the share of final goods revenue in sector $s$ in location $k$ that is spent on tasks performed by location $i$ in occupation $o$ ($\lambda_{kiso} e_{kso}$). To obtain total labor income in location $i$, we sum across sectors $s$, occupations $o$, locations of final goods production $k$ and markets $n$.

Similarly, employment in each sector and location satisfies the equality between payments to workers employed in that sector and location and expenditure on tasks supplied by workers in that sector and location:

$$w_i L_{is} = \sum_o \sum_n \sum_k \lambda_{kiso} e_{kso} \left[ \pi_{nks} E_{ns} w_n L_n \right]. \tag{35}$$

Finally, employment in each occupation, sector and location satisfies the equality between payments to workers employed in that occupation, sector and location and expenditure on tasks supplied by workers in that occupation, sector and location:

$$w_i L_{iso} = \sum_n \sum_k \lambda_{kiso} e_{kso} \left[ \pi_{nks} E_{ns} w_n L_n \right]. \tag{36}$$

## 2.8 General Equilibrium

The general equilibrium of the model can be referenced by the vector of wages for all locations ($w_n$) and the allocation of employment to each occupation, sector and location ($L_{nso}$). Equilibrium wages and employment allocations are determined by the following system of equations:

$$w_i L_i = \sum_{o \in O_s} \sum_{s \in S} \sum_{n \in N} \sum_{k \in N} \lambda_{kiso} e_{kso} \left[ \pi_{nks} E_{ns} w_n L_n \right], \tag{37}$$

$$w_i L_{is} = \sum_{o \in O_s} \sum_{n \in N} \sum_{k \in N} \lambda_{kiso} e_{kso} \left[ \pi_{nks} E_{ns} w_n L_n \right],$$

$$w_i L_{iso} = \sum_{n \in N} \sum_{k \in N} \lambda_{kiso} e_{kso} \left[ \pi_{nks} E_{ns} w_n L_n \right],$$

$$\lambda_{niso} = \frac{U_{iso} L_{iso}^{\chi_{so}} \left( \tau_{niso} w_i \right)^{-\epsilon_{so}}}{\sum_{k \in N} U_{kso} L_{kso}^{\chi_{so}} \left( \tau_{nkso} w_k \right)^{-\epsilon_{so}}}, \tag{38}$$

$$e_{nso} = \frac{\gamma_{so}^{1-\mu_s} \left( \frac{U_{nso} L_{nso}^{\chi_{so}}}{\lambda_{nnso}} \right)^{-\frac{1-\mu_s}{\epsilon_{so}}}}{\sum_{m \in O_s} \gamma_{sm}^{1-\mu_s} \left( \frac{U_{nsm} L_{nsm}^{\chi_{so}}}{\lambda_{nnsm}} \right)^{-\frac{1-\mu_s}{\epsilon_{sm}}}}, \tag{39}$$

$$\pi_{nis} = \frac{T_{is} L_{is}^{\eta_s} \left( d_{nis} \Phi_{is} w_i \right)^{-\theta_s}}{\sum_{k \in N} T_{ks} L_{ks}^{\eta_s} \left( d_{nks} \Phi_{ks} w_k \right)^{-\theta_s}}, \tag{40}$$

$$\Phi_{is} = \left[ \sum_{o \in O_s} \gamma_{so}^{1-\mu_s} \left( \frac{U_{iso} L_{iso}^{\chi_{so}}}{\lambda_{iiso}} \right)^{-\frac{1-\mu_s}{\epsilon_{so}}} \right]^{\frac{1}{1-\mu_s}}, \tag{41}$$

$$E_{ns} = \frac{\kappa_s^{1-\beta} \left( \frac{T_{ns} L_{ns}^{\eta_s}}{\pi_{nns}} \right)^{-\frac{1-\beta}{\theta_s}} \Phi_{ns}^{1-\beta}}{\sum_{r \in S} \kappa_r^{1-\beta} \left( \frac{T_{nr} L_{nr}^{\eta_r}}{\pi_{nnr}} \right)^{-\frac{1-\beta}{\theta_r}} \Phi_{nr}^{1-\beta}}, \tag{42}$$

$$\xi_n = \frac{\left[ \sum_{s \in S} \kappa_s^{-(1-\beta)} \left( \frac{T_{ns} L_{ns}^{\eta_s}}{\pi_{nns}} \right)^{\frac{1-\beta}{\theta_s}} \left[ \sum_{o \in O_s} \gamma_{so}^{-(1-\mu_s)} \left( \frac{U_{nso} L_{nso}^{\chi_{so}}}{\lambda_{nnso}} \right)^{\frac{1-\mu_s}{\epsilon_{so}}} \right]^{\frac{1-\beta}{1-\mu_s}} \right]^{\frac{\alpha}{(1-\alpha)(1-\beta)}} H_n}{\sum_{k \in N} \left[ \sum_{s \in S} \kappa_s^{-(1-\beta)} \left( \frac{T_{ks} L_{ks}^{\eta_s}}{\pi_{kks}} \right)^{\frac{1-\beta}{\theta_s}} \left[ \sum_{o \in O_s} \gamma_{so}^{-(1-\mu_s)} \left( \frac{U_{kso} L_{kso}^{\chi_{so}}}{\lambda_{kkso}} \right)^{\frac{1-\mu_s}{\epsilon_{so}}} \right]^{\frac{1-\beta}{1-\mu_s}} \right]^{\frac{\alpha}{(1-\alpha)(1-\beta)}} H_k}, \tag{43}$$

where $\xi_n = L_n / \bar{L}$ denotes the share of location $n$ in the total population. This system of equations can be solved numerically for arbitrary numbers of locations, sectors and occupations.

### 2.9 Simple Special Case

One simple special case of the model that permits a particularly tractable characterization of general equilibrium is when the following conditions are satisfied: (a) there are no external economies of scale in final goods or task production ($\eta_s = \chi_{so} = 0$) so that productivity in each sector and location is determined solely by exogenous fundamentals $\{T_{is}, U_{iso}\}$, (b) one of the sectors is an outside sector that produces a homogeneous good that is costlessly traded between locations and produced under conditions of perfect competition with a deterministic labor requirement ($Y_{i0} = T_{i0}L_{i0}$ for the outside sector $s = 0$).

In this special case, the model acquires a recursive structure, in which wages can be first determined before determining all the other components of the general equilibrium as a function of wages. We choose the outside good as the numeraire ($p_{i0} = 1$) and consider an equilibrium in which all locations produce the outside good, as can be ensured by the appropriate choice of productivity in this sector for each location. Since the outside good is costlessly traded and produced in all locations, the wage in each location is pinned down by productivity in this sector alone:

$$w_i = T_{i0}.$$

Having determined wages, shares of locations in trade in tasks within all other sectors follow immediately:

$$\lambda_{niso} = \frac{U_{iso}\left(\tau_{niso}w_i\right)^{-\epsilon_{so}}}{\sum_{k\in N} U_{kso}\left(\tau_{nkso}w_k\right)^{-\epsilon_{so}}}, \qquad s \neq 0,$$

from which we obtain the share of occupations in costs for all other sectors:

$$e_{nso} = \frac{\gamma_{so}^{1-\mu_s}\left(\frac{U_{nso}}{\lambda_{nnso}}\right)^{-\frac{1-\mu_s}{\epsilon_{so}}}}{\sum_{m\in O_s}\gamma_{sm}^{1-\mu_s}\left(\frac{U_{nsm}}{\lambda_{nnsm}}\right)^{-\frac{1-\mu_s}{\epsilon_{sm}}}}, \qquad s \neq 0.$$

Having solved for wages and trade in tasks, shares of locations in trade in final goods within each sector follow immediately:

$$\pi_{nis} = \frac{T_{is}\left(d_{nis}\Phi_{is}w_i\right)^{-\theta_s}}{\sum_{k\in N} T_{ks}\left(d_{nks}\Phi_{ks}w_k\right)^{-\theta_s}}, \qquad s \neq 0,$$

$$\Phi_{is} = \left[\sum_{o\in O_s}\gamma_{so}^{1-\mu_s}\left(\frac{U_{iso}}{\lambda_{iiso}}\right)^{-\frac{1-\mu_s}{\epsilon_{so}}}\right]^{\frac{1}{1-\mu_s}}, \qquad s \neq 0,$$

from which we obtain the share of sectors in expenditure:

$$E_{ns} = \frac{\kappa_s^{1-\beta}\left(\frac{T_{ns}}{\pi_{nns}}\right)^{-\frac{1-\beta}{\theta_s}}\Phi_{ns}^{1-\beta}}{1+\sum_{r\neq 0}\kappa_r^{1-\beta}\left(\frac{T_{nr}}{\pi_{nnr}}\right)^{-\frac{1-\beta}{\theta_r}}\Phi_{nr}^{1-\beta}}, \qquad s \neq 0.$$

Finally, having determined trade in tasks and final goods, we obtain population shares:

$$\xi_n = \frac{\left[1+\sum_{s\neq 0}\kappa_s^{-(1-\beta)}\left(\frac{T_{ns}}{\pi_{nns}}\right)^{\frac{1-\beta}{\theta_s}}\left[\sum_{o\in O_s}\gamma_{so}^{-(1-\mu_s)}\left(\frac{U_{nso}}{\lambda_{nnso}}\right)^{\frac{1-\mu_s}{\epsilon_{so}}}\right]^{\frac{1-\beta}{1-\mu_s}}\right]^{\frac{\alpha}{(1-\alpha)(1-\beta)}}\bar{H}_n}{\sum_{k\in N}\left[1+\sum_{s\neq 0}\kappa_s^{-(1-\beta)}\left(\frac{T_{ks}}{\pi_{kks}}\right)^{\frac{1-\beta}{\theta_s}}\left[\sum_{o\in O_s}\gamma_{so}^{-(1-\mu_s)}\left(\frac{U_{kso}}{\lambda_{kkso}}\right)^{\frac{1-\mu_s}{\epsilon_{so}}}\right]^{\frac{1-\beta}{1-\mu_s}}\right]^{\frac{\alpha}{(1-\alpha)(1-\beta)}}\bar{H}_k}.$$

### 2.10 Reductions in Transport and Communication Costs

The distribution of employment across occupations, sectors and locations in the model is determined by two sets of forces: productivity differences (which depend on both an exogenous component and an endogenous component through agglomeration forces) and the costs of trading both tasks and final goods. Together these two sets of forces determine comparative advantages across occupations within sectors and across sectors.

Patterns of comparative advantage across occupations within sectors can be characterized by a double difference for a given import market. The first difference computes the ratio of exports of tasks from two locations $i$ and $k$ in a third market $n$ in a single occupation; the second difference compares this ratio of exports of tasks for two separate occupations $o$ and $m$. Taking this double difference in the unit cost share in equation (17), we obtain:

$$\frac{\lambda_{niso}/\lambda_{nkso}}{\lambda_{nism}/\lambda_{nksm}} = \frac{\left[U_{iso}L_{iso}^{\chi_{so}}\left(\tau_{niso}w_i\right)^{-\epsilon_{so}}\right] / \left[U_{kso}L_{kso}^{\chi_{so}}\left(\tau_{nkso}w_k\right)^{-\epsilon_{so}}\right]}{\left[U_{ism}L_{ism}^{\chi_{so}}\left(\tau_{nism}w_i\right)^{-\epsilon_{sm}}\right] / \left[U_{ksm}L_{ksm}^{\chi_{so}}\left(\tau_{nksm}w_k\right)^{-\epsilon_{sm}}\right]}. \tag{44}$$

Therefore, a location $i$ specializes more in occupation $o$ relative to occupation $m$ compared to another location $k$ when it has lower production costs (as determined by wages $w_i$, the exogenous productivity parameter $U_{iso}$, and the endogenous component of productivity from agglomeration forces $L_{iso}^{\chi_{so}}$) and lower bilateral costs of trading tasks (as determined by $\tau_{niso}$)

Patterns of comparative advantage across sectors can be characterized by an analogous double difference for a given import market. The first difference computes the ratio of exports of final goods from two locations $i$ and $k$ in a third market $n$ in a single sector; the second difference compares this ratio of exports of final goods for two separate sectors $s$ and $r$. Taking this double difference in the expenditure share in equation (25), we obtain:

$$\frac{\pi_{nis}/\pi_{nks}}{\pi_{nir}/\pi_{nkr}} = \frac{\left[T_{is}L_{is}^{\eta_s}\left(d_{nis}\Phi_{is}w_i\right)^{-\theta_s}\right] / \left[T_{ks}L_{ks}^{\eta_s}\left(d_{nks}\Phi_{ks}w_k\right)^{-\theta_s}\right]}{\left[T_{ir}L_{ir}^{\eta_r}\left(d_{nir}\Phi_{ir}w_i\right)^{-\theta_r}\right] / \left[T_{kr}L_{kr}^{\eta_r}\left(d_{nkr}\Phi_{kr}w_k\right)^{-\theta_r}\right]}. \tag{45}$$

Therefore a location $i$ specializes more in sector $s$ relative to sector $r$ compared to another location $k$ when it has lower production costs (as determined by wages $w_i$, the unit cost summary statistic $\Phi_{is}$, the exogenous productivity parameter $T_{is}$, and the endogenous component of productivity from agglomeration forces $L_{is}^{\eta_s}$) and lower bilateral costs of trading final goods (as determined by $d_{nis}$).

When the costs of trading tasks and final goods are large, all locations have similar employment structures across sectors, and all tasks within each sector are undertaken in the same location where the final good is produced. As the costs of trading final goods and tasks fall, locations specialize across sectors and across occupations within sectors according to their comparative advantage as determined by productivity differences. If densely-populated urban locations have a comparative advantage in interactive tasks relative to sparsely-populated rural locations (e.g. Gaspar and Glaeser 1998), the model predicts that a fall in the costs of trading tasks leads to an increase in the interactiveness of employment within sectors in urban relative to rural areas.

## References


[1] Autor, David H., Frank Levy and Richard J. Murnane (2003) "The Skill Content of Recent Technological Change: An Empirical Exploration," *Quarterly Journal of Economics*, November, 1279-1333.

[2] Costinot, Arnaud, Dave Donaldson and Ivana Komunjer (2012) "What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas," *Review of Economic Studies*, 79(2), 581-608.

[3] Davis, Morris A. and François Ortalo-Magne (2011) "Household Expenditures, Wages, Rents," *Review of Economic Dynamics*, 14(2), 248-261.

[4] Department of Transportation (1976) *America's Highways 1776-1976: A History of the Federal Aid Program*, Washington, U.S. Government Printing Office

[5] Eaton, Jonathan and Samuel S. Kortum (2002) "Technology, Geography, and Trade," *Econometrica*, 70(5) (2002), 1741-1779.

[6] Ethier, Wilfred J. (1982) "Decreasing Costs in International Trade and Frank Grahams Argument for Protection," *Econometrica*, 50, 1243-1268.

[7] Gaspar, Jess and Edward L. Glaeser (1998) "Information Technology and the Future of Cities," *Journal of Urban Economics*, 43(1), 136-156.

[8] Grossman, Gene and Esteban Rossi-Hansberg (2008) "Trading Tasks: A Simple Theory of Offshoring," *American Economic Review*, 98(5), 1978-1997.

[9] Grossman, Gene and Esteban Rossi-Hansberg (2012) "Task Trade between Similar Countries," *Econometrica*, 80(2), 593-629.

[10] Helpman, Elhanan (1998) "The Size of Regions," In *Topics in Public Economics: Theoretical and Applied Analysis*, ed. David Pines, Efraim Sadka, and Itzhak Zilcha, Cambridge: Cambridge University Press.

[11] Ngai, Rachel, and Chris Pissarides (2007) "Structural Change in a Multi-sector Model of Growth," *American Economic Review*, 97(1), 429-443.

[12] Yi, Kei-Mu and Jing Zhang (2013) "Structural Change in an Open Economy," *Journal of Monetary Economics*, 60(6), 667-682.

**Table A1: Specialization Across Narrow Occupations 1880-2000**

| Panel A: Top 20 occupations whose concentration in metro areas increased most from 1880-2000 | Difference in Ranks 2000-1880 | Panel B: Top 20 occupations whose concentration in metro areas increased least from 1880-2000 | Difference in Ranks 2000-1880 |
|---|---|---|---|
| Geologists and geophysicists | -141 | Tool makers, and die makers and setters | 58 |
| Lawyers and judges | -130 | Bookbinders | 59 |
| Physicians and surgeons | -128 | Inspectors (nec) | 61 |
| Members of the armed services | -121 | Stationary firemen | 63 |
| Biological scientists | -119 | Newsboys | 71 |
| Mining-Engineers | -118 | Janitors and sextons | 74 |
| Dentists | -114 | Recreation and group workers | 75 |
| Plasterers | -112 | Conductors, railroad | 76 |
| Officials and administratators (nec), public administration | -106 | Dispatchers and starters, vehicle | 78 |
| Editors and reporters | -99 | Oilers and greaser, except auto | 79 |
| Buyers and dept heads, store | -93 | Filers, grinders, and polishers, metal | 81 |
| Civil-Engineers | -88 | Sawyers | 82 |
| Professional, technical and kindred workers (nec) | -83 | Taxicab drivers and chauffeurs | 83 |
| Dancers and dancing teachers | -82 | Painters, except construction or maintenance | 84 |
| Painters, construction and maintenance | -75 | Charwomen and cleaners | 91 |
| Tinsmiths, coppersmiths, and sheet metal workers | -74 | Welders and flame cutters | 92 |
| Managers and superintendants, building | -74 | Cement and concrete finishers | 101 |
| Teachers (n.e.c.) | -73 | Weavers, textile | 105 |
| Pattern and model makers, except paper | -71 | Upholsterers | 126 |
| Jewelers, watchmakers, goldsmiths, and silversmiths | -68 | Veterinarians | 131 |

Notes: Coefficients estimated from a regression of a (0,1) dummy variable for whether a worker is located in a metro area on occupation-year and sector-year fixed effects (regression (1) in the paper). Occupation-year and sector-year fixed effects are each normalized to sum to zero. A separate regression is estimated for each year. Standard errors are clustered by occupation.

**Table A2: Verbs Most and Least Strongly Correlated with Metro Area Employment Shares (1939 DOTs)**

| Panel A: Verbs Most Strongly Correlated with Metro Area Employment Shares | | | | | | | |
|---|---|---|---|---|---|---|---|
| Rank | 1880 | 1900 | 1920 | 1940 | 1960 | 1980 | 2000 |
| 1 | Retouch | Permeate | Permeate | Bounce | Accrue | Estimate | Advise |
| 2 | Flounce | Flounce | Fake | Blaze | Kid | Advise | Sell |
| 3 | Permeate | Retouch | Hum | Reserve | Undercharge | Calculate | Estimate |
| 4 | Lure | Initiate | Seep | Converge | Seep | Appraise | Investigate |
| 5 | Abut | Report | Kid | Favor | Prompt | Investigate | Prefer |
| 6 | Highlight | Enamel | Undercharge | Mail | Converge | Question | Appraise |
| 7 | Solidify | Solidify | Smash | Kid | Necessitate | Accrue | Quote |
| 8 | Glow | Refund | Accrue | Undercharge | Document | Wage | Display |
| 9 | Trawl | Enlarge | Necessitate | Lobby | Allocate | Adjudicate | Bid |
| 10 | Finish | Identify | Overload | Seep | Doff | Inform | Jail |
| Panel A: Verbs Least Strongly Correlated with Metro Area Employment Shares | | | | | | | |
| Rank | 1880 | 1900 | 1920 | 1940 | 1960 | 1980 | 2000 |
| 1682 | Dovetail | Lift | Demand | Program | Lounge | Flout | Narrow |
| 1683 | Sail | Rain | Discuss | Transact | Spoon | Heft | Line |
| 1684 | Overturn | Pink | Resist | Seam | Encounter | Sinter | Transport |
| 1685 | Extend | Finish | Induce | Grapple | Sinter | Hang | Truck |
| 1686 | Attain | Sew | Spoon | Board | Smoke | Hook | Drive |
| 1687 | Late | Top | Snarl | Pick | Back | Truck | Remove |
| 1688 | Embroider | Notch | Resume | Back | Pile | Bolt | Screw |
| 1689 | Fudge | Embroider | Intersperse | Flicker | Hand | Remove | Hook |
| 1690 | Foul | Offset | Top | Resume | Boat | Line | Bolt |
| 1691 | Offset | Scrutinize | Recede | Recede | Anchor | Hand | Hand |

Notes: Coefficients estimated from a regression of the share of occupation-sector employment in metro areas on the frequency with which a verb is used for an occupation and sector-year fixed effects (regression (2) in the paper). A separate regression is estimated for each verb. Verbs are sorted by their estimated coefficients normalized by the standard deviation for the verb frequency. Verbs are from the time-invariant occupational descriptions from the 1939 Dictionary of Occupations (DOTs).

**Table A3: Correlations Between Occupational Characteristics**

| Panel A: Unweighted Correlations | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Interactiveness | Interactiveness 1939 | Nonroutine analytic (math) | Nonroutine interactive (dcp) | Routine cognitive (sts) | Routine manual (finger) | Nonroutine manual (ehf) |
| Interactiveness | 1 | | | | | | |
| Interactiveness 1939 | 0.62*** | 1 | | | | | |
| Nonroutine analytic (math) | 0.55*** | 0.48*** | 1 | | | | |
| Nonroutine interactive (dcp) | 0.47*** | 0.44*** | 0.54*** | 1 | | | |
| Routine cognitive (sts) | -0.33*** | -0.27*** | 0.25*** | -0.18** | 1 | | |
| Routine manual (finger) | -0.09 | -0.11 | 0.27*** | -0.08 | 0.52*** | 1 | |
| Nonroutine manual (ehf) | -0.39*** | -0.19** | -0.31*** | -0.17** | 0.003 | -0.09 | 1 |

| Panel B: Weighted Correlations | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Interactiveness | Interactiveness 1939 | Nonroutine analytic (math) | Nonroutine interactive (dcp) | Routine cognitive (sts) | Routine manual (finger) | Nonroutine manual (ehf) |
| Interactiveness | 1 | | | | | | |
| Interactiveness 1939 | 0.52*** | 1 | | | | | |
| Nonroutine analytic (math) | 0.54*** | 0.32*** | 1 | | | | |
| Nonroutine interactive (dcp) | 0.47*** | 0.12 | 0.59*** | 1 | | | |
| Routine cognitive (sts) | -0.25*** | -0.03 | 0.22*** | -0.33*** | 1 | | |
| Routine manual (finger) | -0.14* | -0.04 | 0.07 | -0.30*** | 0.38*** | 1 | |
| Nonroutine manual (ehf) | -0.47*** | -0.22*** | -0.31*** | -0.21*** | 0.05 | -0.09 | 1 |

Note: Table reports correlations between occupational characteristics across the sample of occupations in 2000. Interactiveness is our baseline measure using occupational descriptions from the 1991 DOTs. Interactiveness 1939 is our robustness measure using occupational descriptions from the 1939 DOTs. Nonroutine analytic, Nonroutine interactive, Routine cognitive, Routine manual and Nonroutine manual are measures based on the numerical scores in the 1991 DOTs as used in Autor, Levy and Murnane (2003). Weighted correlations are weighted by occupation employment in 2000. * significant at 10%; ** significant at 5%; *** significant at 1%.

**Table A4: Top Five Verbs for Each Thesaurus Section**

| Thesaurus Section | Verb1 | Verb2 | Verb3 | Verb4 | Verb5 |
|---|---|---|---|---|---|
| 1 Class I. Words Expressing Abstract Relations | | | | | |
| 1.1 Section I. Existence | Exist | Zero | Obtain | Posture | Continue |
| 1.2 Section II. Relation | Adapt | Correspond | Reprint | Contrast | Harmonize |
| 1.3 Section III. Quantity | Solder | Cleave | Blend | Clamp | Latch |
| 1.4 Section IV. Order | Sample | Specialize | Include | Disperse | Cluster |
| 1.5 Section V. Number | Recur | Halve | Invoice | Schedule | Compute |
| 1.6 Section VI. Time | Date | Modernize | Synchronize | Dial | Late |
| 1.7 Section VII. Change | Undergo | Transform | Vary | Change | Arrive |
| 1.8 Section VIII. Causation | Generate | Sheathe | Energize | Fertilize | Heir |
| 2 Class II. Words Relating to Space | | | | | |
| 2.1 Section I. Space in General | Pouch | Reside | Locate | Camp | Vacate |
| 2.2 Section II. Dimensions | Mesh | Widen | Bunk | Tape | Flake |
| 2.3 Section III. Form | Curl | Spike | Envelope | Gouge | Scoop |
| 2.4 Section IV. Motion | Bob | Shunt | Ramp | Dive | Export |
| 3 Class III. Words Relating to Matter | | | | | |
| 3.1 Section I. Matter in General | Weigh | Float | Swim | Balloon | Pound |
| 3.2 Section II. Inorganic Matter | Grease | Irrigate | Soak | Liquefy | Lard |
| 3.3 Section III. Organic Matter | Tint | Glare | Smell | Chime | Bleach |
| 4 Class IV. Words Relating to the Intellectual Faculties | | | | | |
| 4.1 Division I. Formation of Ideas | | | | | |
| 4.1.1 Section I. Intellect in General | Occur | Discuss | Weigh | Loop | Consider |
| 4.1.2 Section II. Precursory Conditions and Operations | Assay | Examine | Scrutinize | Experiment | Trawl |
| 4.1.3 Section III. Materials for Reasoning | Ensure | Attest | Authenticate | Testify | Insure |
| 4.1.4 Section IV. Reasoning Processes | Disprove | Guess | Defeat | Demonstrate | Mystify |
| 4.1.5 Section V. Results Of Reasoning | Conform | Minimize | Adjudicate | Detect | Unlock |
| 4.1.6 Section VI. Extension of Thought | Predict | Memorize | Forecast | Announce | Anticipate |
| 4.1.7 Section VII. Creative Thought | Visualize | Guess | Create | Devise | Fabricate |
| 4.2 Division II. Communication of Ideas | | | | | |
| 4.2.1 Section I. Nature of Ideas Communicated | Annotate | Decipher | Interpret | Fudge | Clarify |
| 4.2.2 Section II. Modes of Communication | Disguise | Fake | Learn | Educate | Teach |
| 4.2.3 Section III. Means of Communicating Ideas | Write | Describe | Narrate | Relate | Underlay |
| 5 Class V. Words Relating to the Voluntary Powers | | | | | |
| 5.1 Division I. Individual Volition | | | | | |
| 5.1.1 Section I. Volition in General | Familiarize | Incline | Volunteer | Deflate | Deter |
| 5.1.2 Section II. Prospective Volition | Rot | Drug | Poison | Purify | Misuse |
| 5.1.3 Section III. Voluntary Action | Manage | Consult | Fatigue | Transact | Confer |
| 5.1.4 Section IV. Antagonism | Contest | Bombard | Assist | Avert | Obstruct |
| 5.1.5 Section V. Results of Voluntary Action | Abort | Accomplish | Defeat | Drown | Blossom |
| 5.2 Division II. Social Volition | | | | | |
| 5.2.1 Section I. General Intersocial Volition | Restrain | Liberate | Ballot | Delegate | Curb |
| 5.2.2 Section II. Special Intersocial Volition | Petition | Prohibit | Authorize | Permit | Invite |
| 5.2.3 Section III. Conditional Intersocial Volition | Underwrite | Pawn | Endorse | Observe | Insure |
| 5.2.4 Section IV. Possessive Relations | Afford | Finance | Liquidate | Grab | Clutch |
| 6 Class VI. Emotion, Religion and Morality | | | | | |
| 6.1 Section I. Affections in General | Awaken | Animate | Excite | Impress | Stipulate |
| 6.2 Section II. Personal Affections | Enliven | Fear | Reassure | Beautify | Decorate |
| 6.3 Section III. Sympathetic Affections | Snarl | Welcome | Kiss | Visit | Butcher |
| 6.4 Section IV. Moral Affections | Switch | Thresh | Police | Tipple | Disapprove |
| 6.5 Section V. Religious Affections | Anoint | Induct | Translate | Justify | Cure |

Note: Verbs most concentrated in each thesaurus section (verbs with the top five values of $ThesFreq_{vk}$ from equation (3) in the paper for each thesaurus section, where verb 1 is the highest ranked). Verbs are first sorted by their number of occurrences in a thesaurus section divided by their total number of occurrences in the Dictionary of Occupational Titles (DOTs) for 1991. If two or more verbs have the same value of this fraction, they are next sorted by their number of occurrences in the DOTs, and then next sorted by their alphabetical order.

**Table A5: Correlations with Independent Measures of Interactiveness**

| Unweighted Correlations | | Weighted Correlations | |
|---|---|---|---|
| | Interactiveness | | Interactiveness |
| Interactiveness | 1 | Interactiveness | 1 |
| Assisting and caring for others | 0.22*** | Assisting and caring for others | 0.18** |
| Coaching and developing others | 0.43*** | Coaching and developing others | 0.27*** |
| Communicating with persons outside organization | 0.65*** | Communicating with persons outside organization | 0.63*** |
| Communicating with Supervisors, Peers, or Subordinates | 0.51*** | Communicating with Supervisors, Peers, or Subordinates | 0.53*** |
| Coordinating the work and activities of others | 0.36*** | Coordinating the work and activities of others | 0.41*** |
| Developing and building teams | 0.43*** | Developing and building teams | 0.40*** |
| Establishing and maintaining interpersonal relationships | 0.66*** | Establishing and maintaining interpersonal relationships | 0.70*** |
| Guiding, directing and motivating subordinates | 0.38*** | Guiding, directing and motivating subordinates | 0.38*** |
| Interpreting the meaning of information for others | 0.55*** | Interpreting the meaning of information for others | 0.47*** |
| Monitoring and controlling resources | 0.36*** | Monitoring and controlling resources | 0.21*** |
| Performing administrative activities | 0.70*** | Performing administrative activities | 0.65*** |
| Performing for or working directly with the public | 0.33*** | Performing for or working directly with the public | 0.38*** |
| Resolving conflict and negotiating with others | 0.61*** | Resolving conflict and negotiating with others | 0.66*** |
| Provide consultation and advice to others | 0.56*** | Provide consultation and advice to others | 0.52*** |
| Selling or influencing others | 0.45*** | Selling or influencing others | 0.08 |
| Staffing organizational units | 0.50*** | Staffing organizational units | 0.52*** |
| Training and teaching others | 0.40*** | Training and teaching others | 0.32*** |

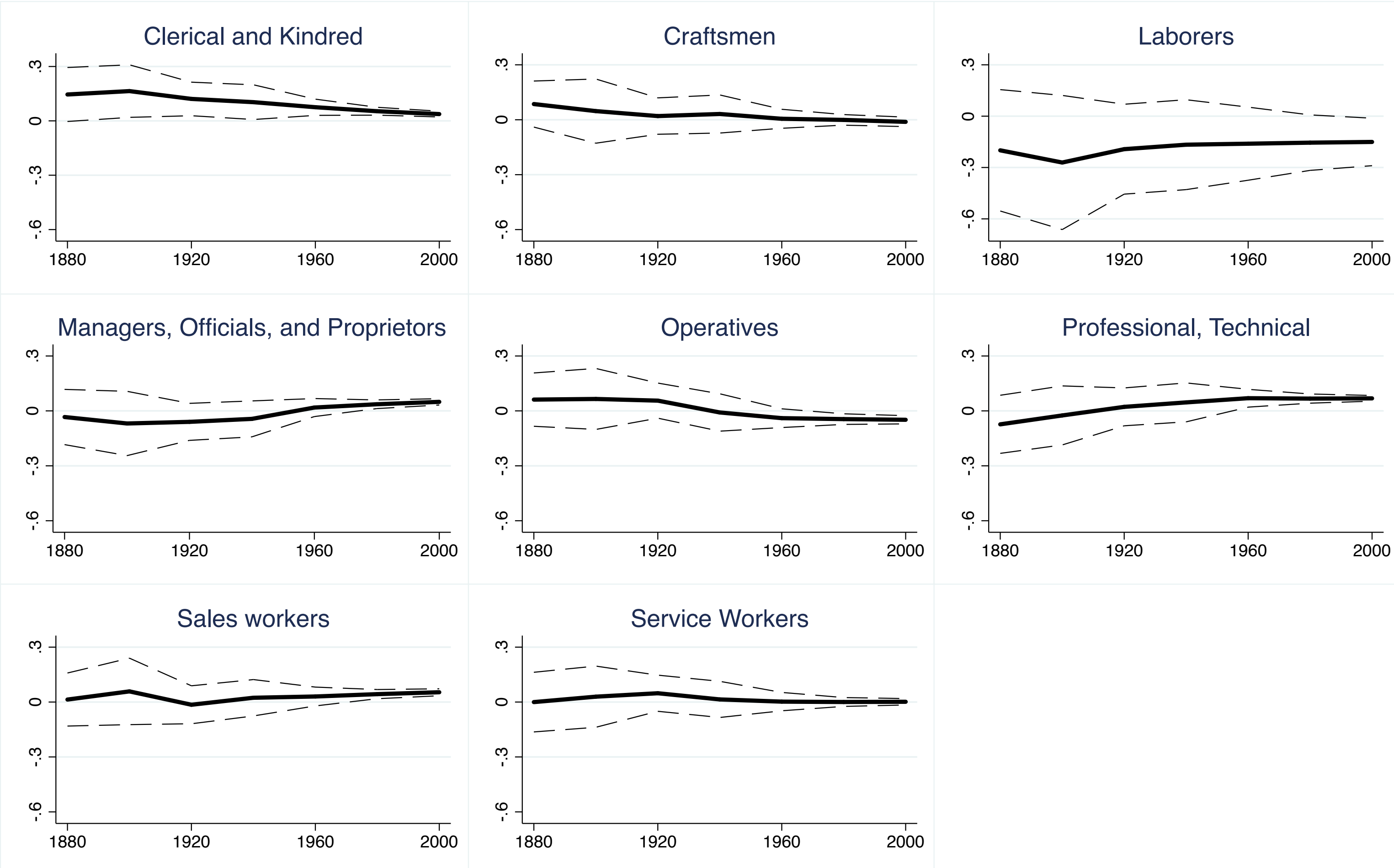
Note: Table reports correlations across occupations between our measure of occupation interactiveness based on the frequency with which verbs from time-invariant occupational descriptions in the 1991 DOTs appear in Class IV, Division 1 (Formation of Ideas), Class IV, Division 2 (Communication of Ideas) and Class V, Division 2 (Intersocial Volition) 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.

**Table A6: Metro Employment and Wagebill Shares and Interactiveness**

| Panel A: Between sectors | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **LHS** | **Measure** | **1880** | **1900** | **1920** | **1940** | **1960** | **1980** | **2000** |
| Employment | Interactiveness | -0.038 | -0.044 | 0.071 | 0.134 | 0.200*** | 0.266*** | 0.207*** |
| Employment | Thought | -0.224*** | -0.400*** | -0.452*** | -0.168 | 0.053 | 0.213*** | 0.305*** |
| Employment | Communication | -0.209*** | -0.275*** | -0.260*** | -0.097 | 0.100 | 0.165* | 0.207** |
| Employment | Intersocial | -0.193** | -0.281*** | -0.306*** | -0.064 | 0.062 | 0.161** | 0.182*** |
| Employment | Individual volition | -0.106*** | -0.152*** | -0.189*** | -0.137*** | -0.088** | 0.015 | 0.056 |
| Wagebill | Interactiveness | | | | 0.135 | 0.148* | 0.235*** | 0.183*** |
| **Panel B: Within sectors** | | | | | | | | |
| **LHS** | **Measure** | **1880** | **1900** | **1920** | **1940** | **1960** | **1980** | **2000** |
| Employment | Interactiveness | -0.088 | -0.058 | -0.026 | -0.009 | 0.053*** | 0.092*** | 0.103*** |
| Employment | Thought | -0.062*** | -0.068*** | -0.059*** | -0.028 | 0.033*** | 0.055*** | 0.060*** |
| Employment | Communication | -0.011*** | -0.010*** | 0.007 | 0.032 | 0.53*** | 0.053*** | 0.044*** |
| Employment | Intersocial | -0.009** | -0.022*** | -0.005 | 0.005 | 0.033*** | 0.022* | 0.016 |
| Employment | Individual volition | -0.064*** | -0.041*** | -0.018 | -0.013 | 0.006 | 0.016 | 0.030** |
| Wagebill | Interactiveness | | | | 0.011 | 0.056*** | 0.091*** | 0.098*** |
| | Sector-year fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

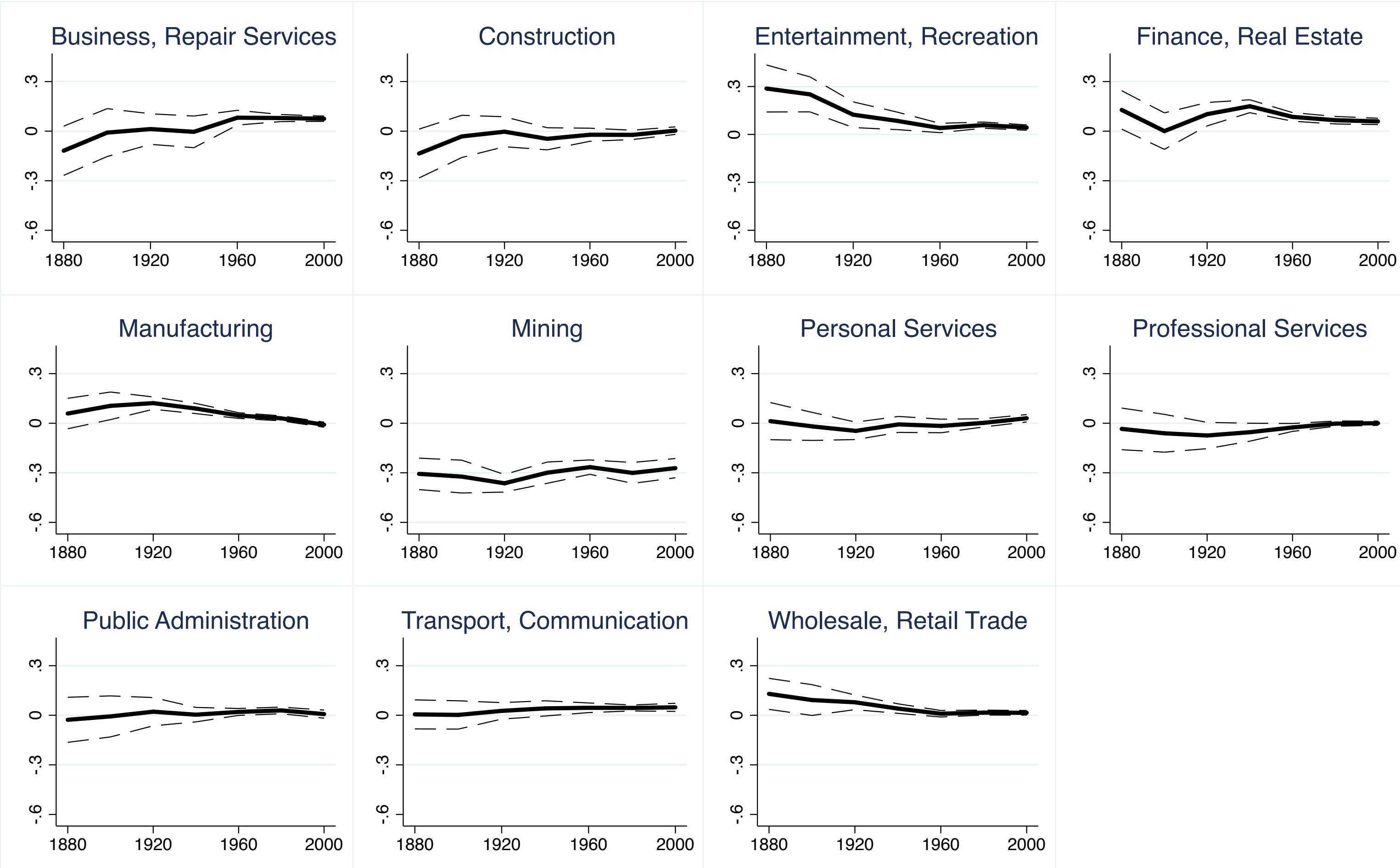
Notes: This table reports beta coefficients for the regression results from in Table 5 in the paper: the estimated coefficients in Table 5 are multiplied by the standard deviation of the independent variable and divided by the standard deviation of the dependent variable. Each cell of each panel of the table corresponds to a separate regression. Coefficients estimated from a regression of the share of either employment or the wagebill in metro areas on the frequency with which the verbs from occupational descriptions appear in a thesaurus section; the wagebill data are only available from 1940 onwards; the frequency with which verbs appear in a thesaurus section is measured using time-invariant occupational descriptions from the 1991 Dictionary of Occupations (DOTs); Interactiveness is the frequency with which verbs from occupational descriptions appear in Class IV, Division 1 (Formation of Ideas), Class IV, Division 2 (Communication of Ideas) and Class V, Division 2 (Intersocial Volition) of the thesaurus; Thought is the frequency with which verbs appear in Class IV (Division 1) of the thesaurus; Communication is the frequency with which verbs appear in Class IV (Division 2) of the thesaurus; Intersocial is the frequency with which verbs appear in Class V (Division 2) of the thesaurus; Individual volition is the frequency with which verbs appear in Class V (Division 1). In Panel A, observations are three-digit sectors for each year, the frequency of verb use for each sector is the employment-weighted average of the frequency for occupations within that sector, and the standard errors reported in Table 5 in the paper are heteroskedasticity robust (equation (27) in the paper). In Panel B, observations are three-digit sectors and occupations for each year, three-digit sector fixed effects are included, and the standard errors reported in Table 5 in the paper are heteroskedasticity robust and clustered on occupation (equation (28) in the paper). * significant at 10%; ** significant at 5%; *** significant at 1%.

# Figure A1: Metro Area Specialization for Aggregate Occupations



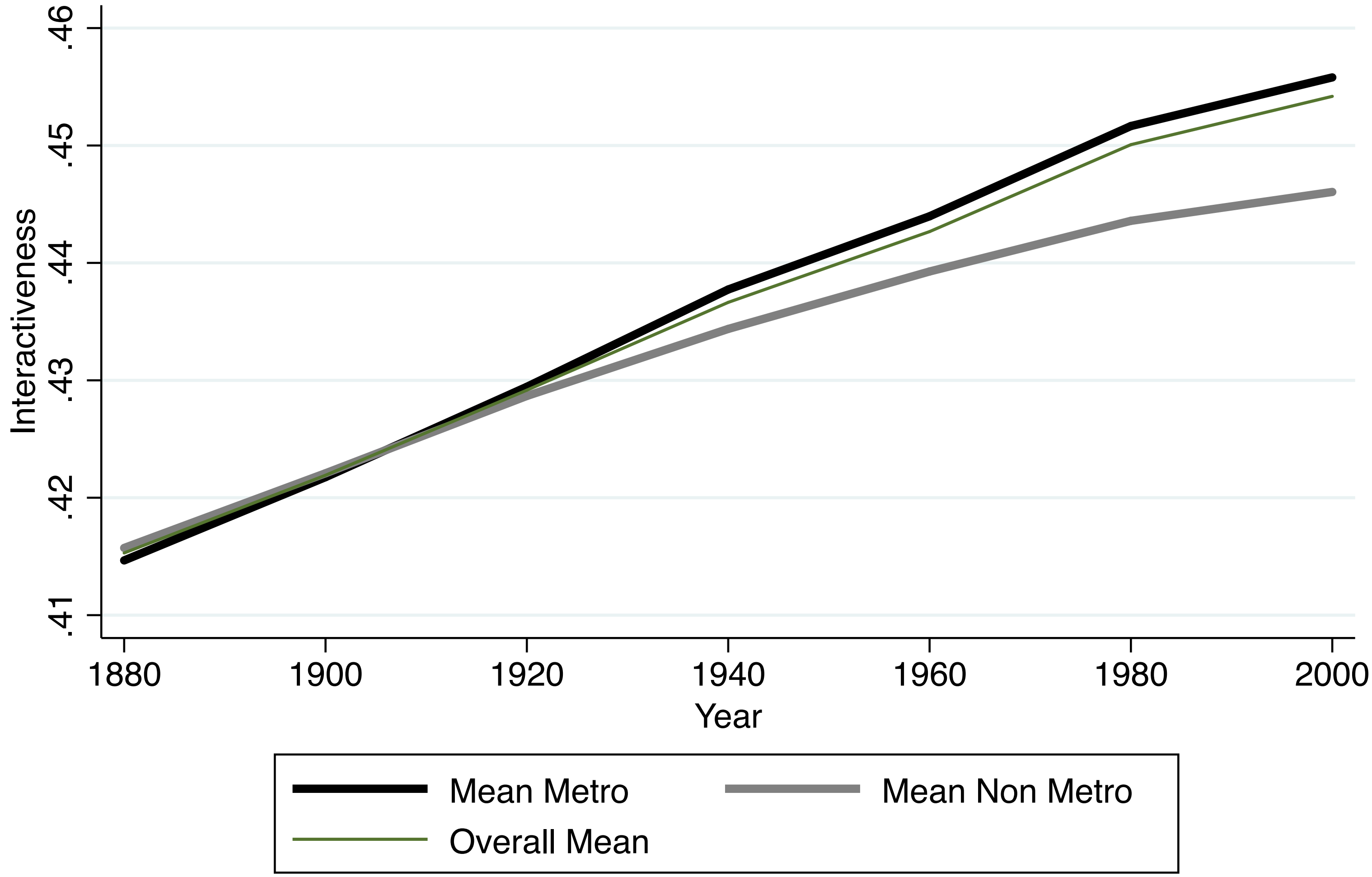

Notes: Coefficients estimated from a regression of an indicator variable for whether a worker is located in a metro area on occupation-year and sector-year fixed effects (equation (1) in the paper). Occupation-year and sector-year fixed effects are each normalized to sum to zero in each year. A separate regression is estimated for each year.

# Figure A2: Metro Area Specialization for Aggregate Sectors



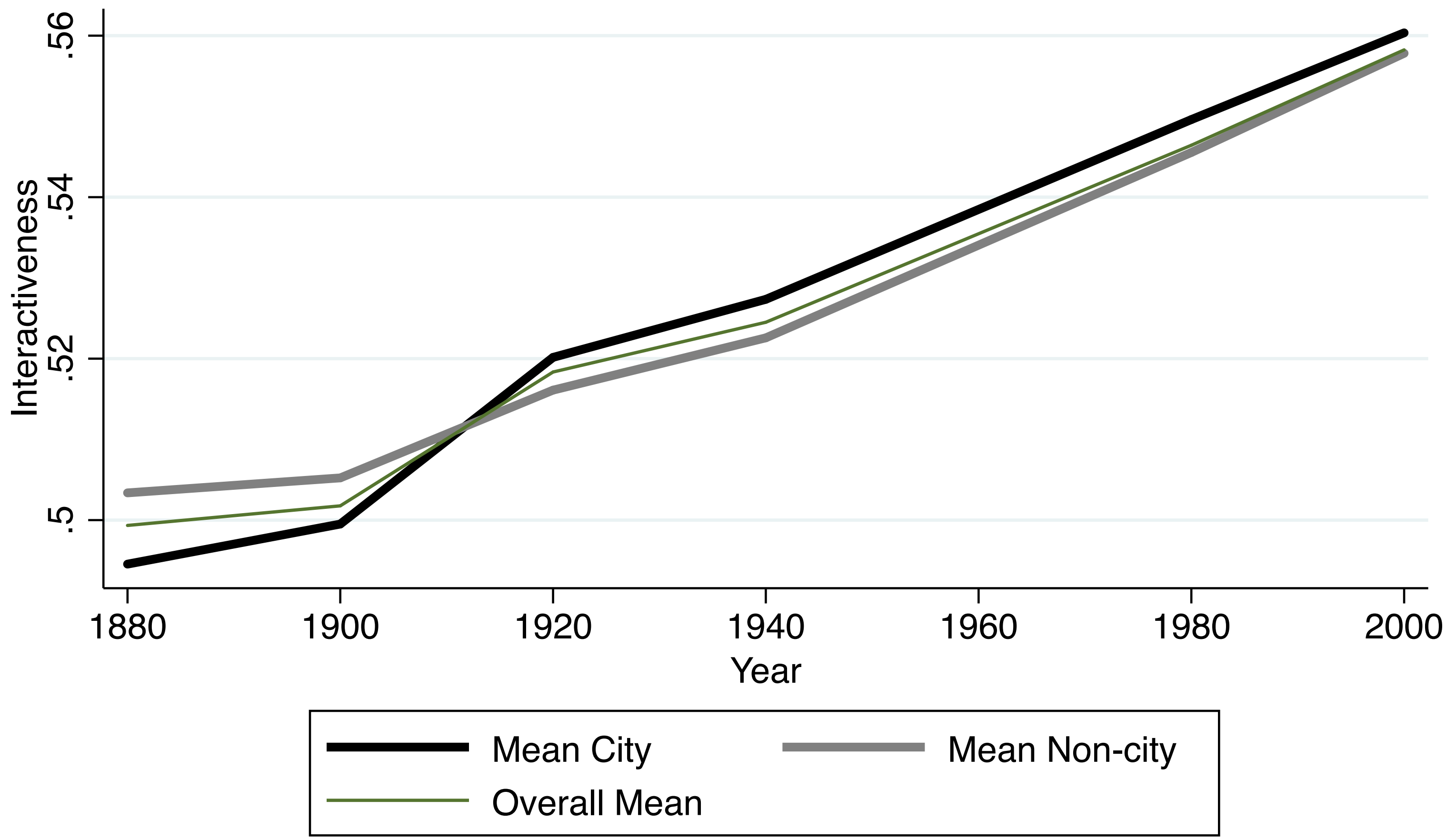

Notes: Coefficients estimated from a regression of an indicator variable for whether a worker is located in a metro area on occupation-year and sector-year fixed effects (regression (11) in the paper). Occupation-year and sector-year fixed effects are each normalized to sum to zero. A separate regression is estimated for each year.

## Figure A3: Mean 1939 Interactiveness in Metro and Non-Metro Areas



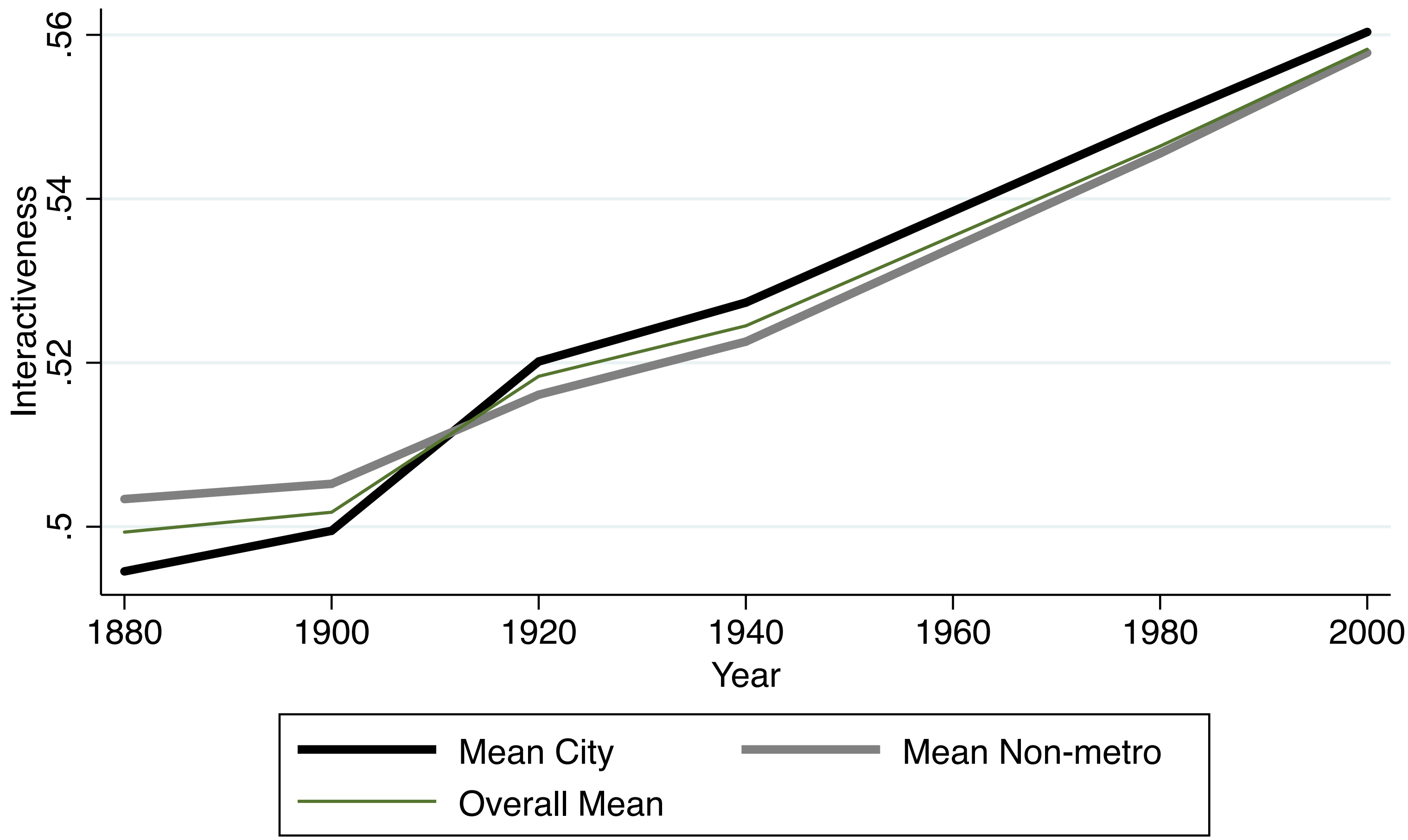

Note: Mean interactiveness computed using time-invariant occupational descriptions from the 1939 DOTs.

Figure A4: Mean Interactiveness in Administrative Cities versus All Other Areas



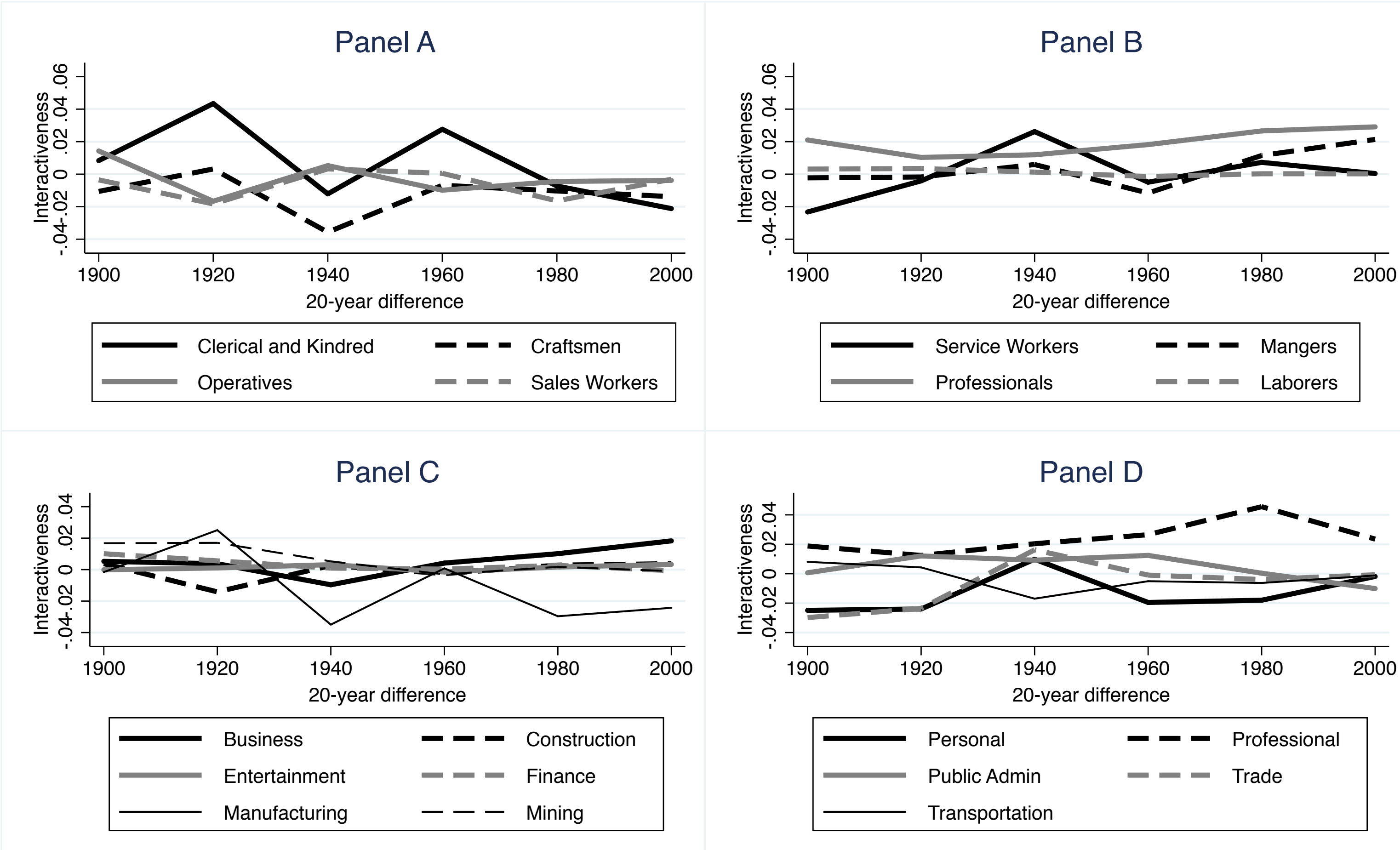

Note: Mean interactiveness computed using time-invariant occupational descriptions from the 1991 DOTs. Non-administrative cities includes both non-metro areas and the parts of metro areas outside administrative cities. The administrative cities indicator is not available in 1960 in IPUMs and hence 1960 is omitted from the figure.

Figure A5: Mean Interactiveness in Administrative Cities versus Non-Metro Areas



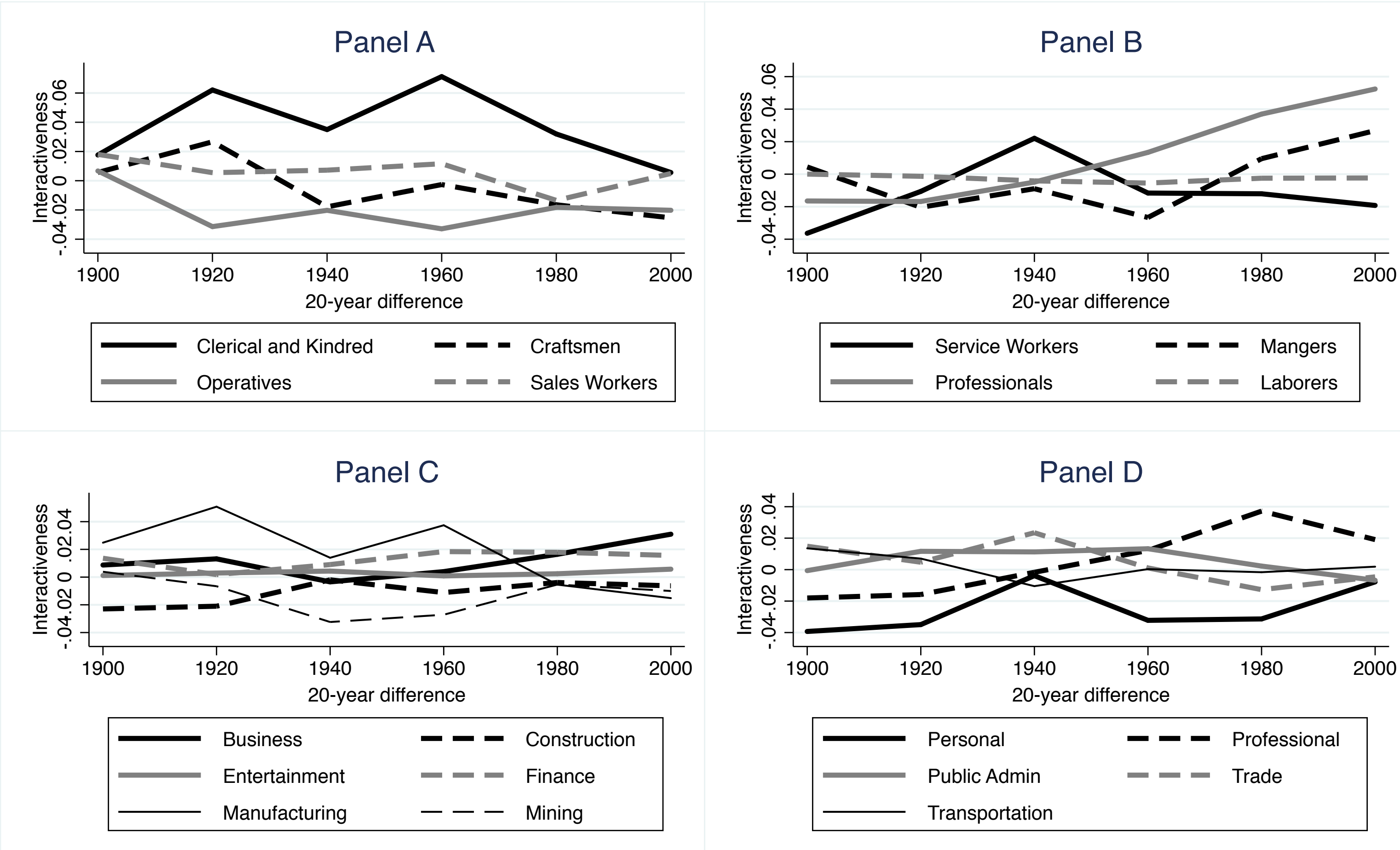

Note: Mean interactiveness computed using time-invariant occupational descriptions from the 1991 DOTs. The administrative cities indicator is not available in 1960 in IPUMs and hence 1960 is omitted from the figure.

## Figure A6: Decomposition of the Change in Mean Interactiveness in Metro Areas



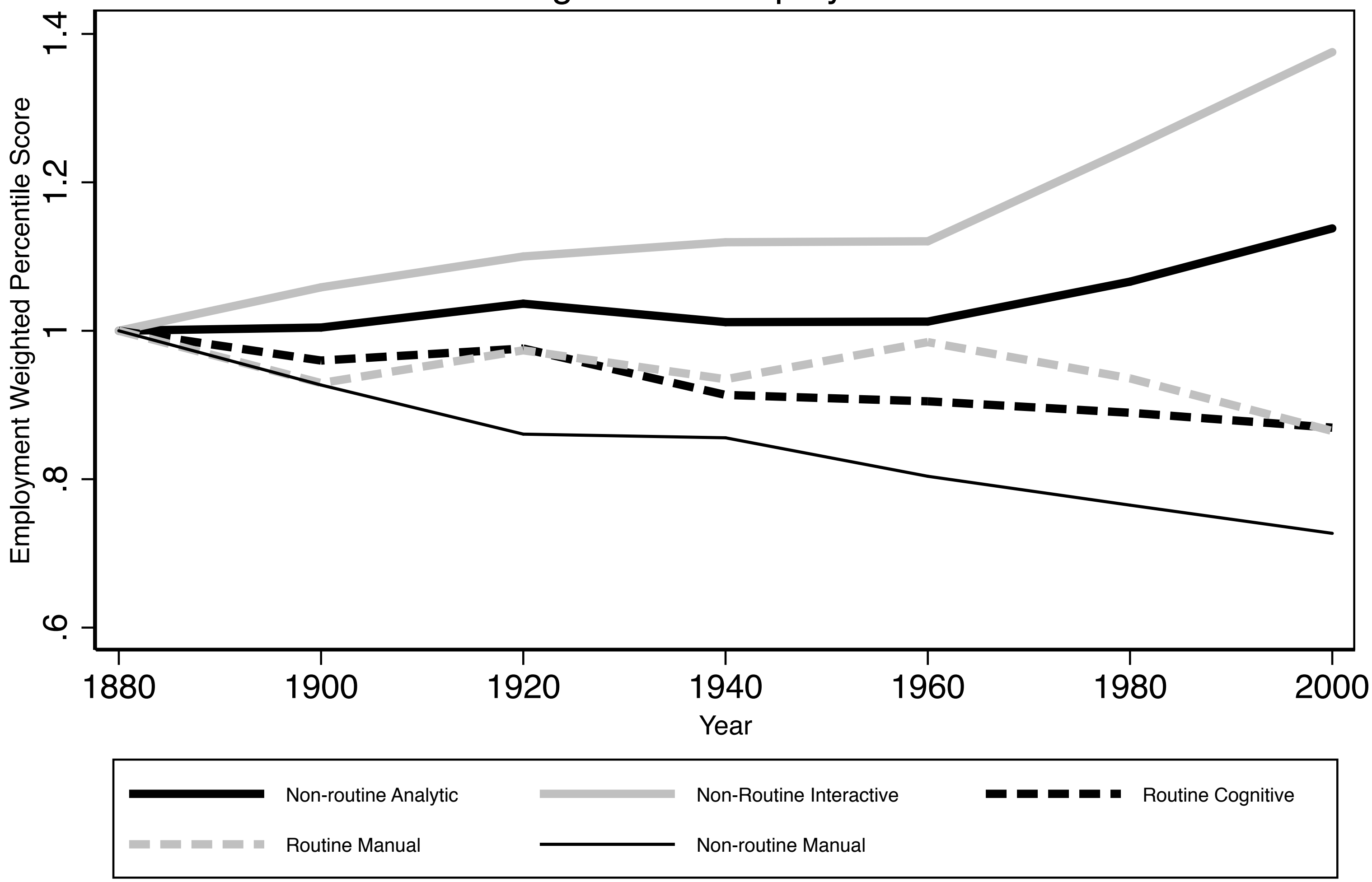

Notes: Decomposition of the change in mean interactiveness in metro areas (equation (25) in the paper) into the contributions of two-digit occupations and sectors. Mean interactiveness based on time-invariant occupational descriptions from the 1991 DOTs.

## Figure A7: Decomposition of the Change in Mean Interactiveness in Non-Metro Areas



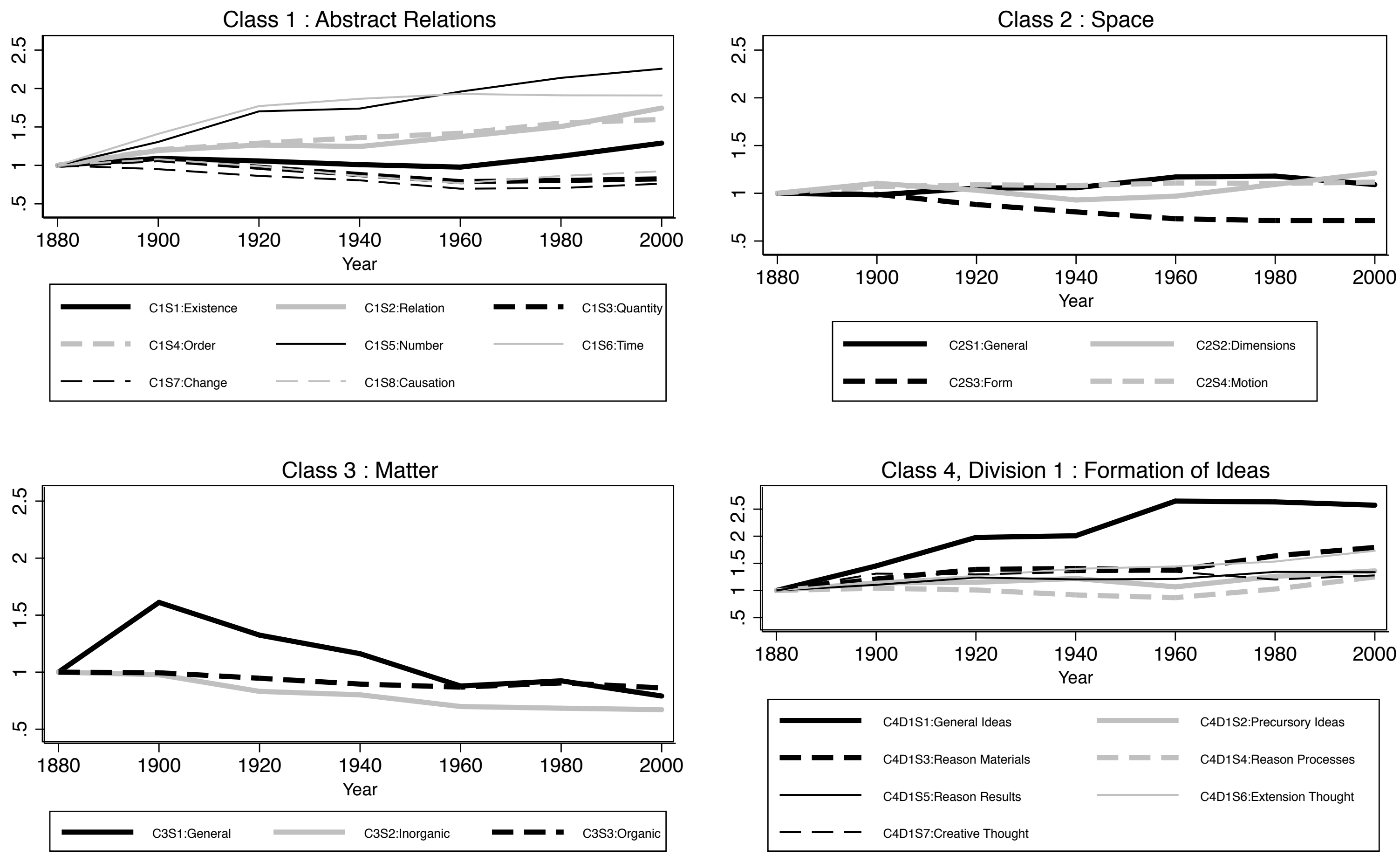

Notes: Decomposition of the change in mean interactiveness in metro areas (equation (25) in the paper) into the contributions of two-digit occupations and sectors. Mean interactiveness based on time-invariant occupational descriptions from the 1991 DOTs.

## Figure A8: Employment



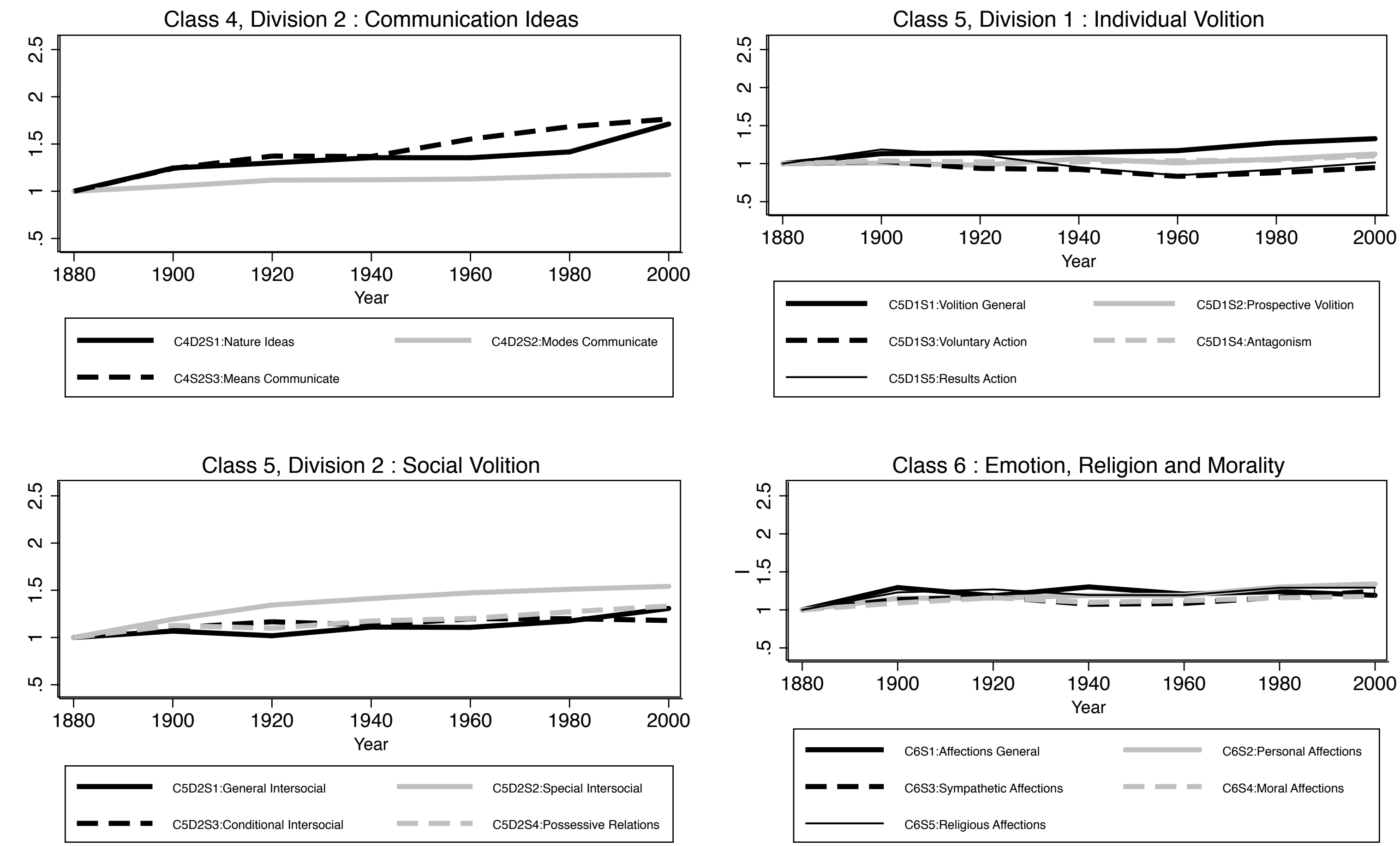

Note: Employment-weighted mean of 1991 DOTs numerical scores in each year, as used in Autor, Levy and Murnane (2003).

# Figure A9



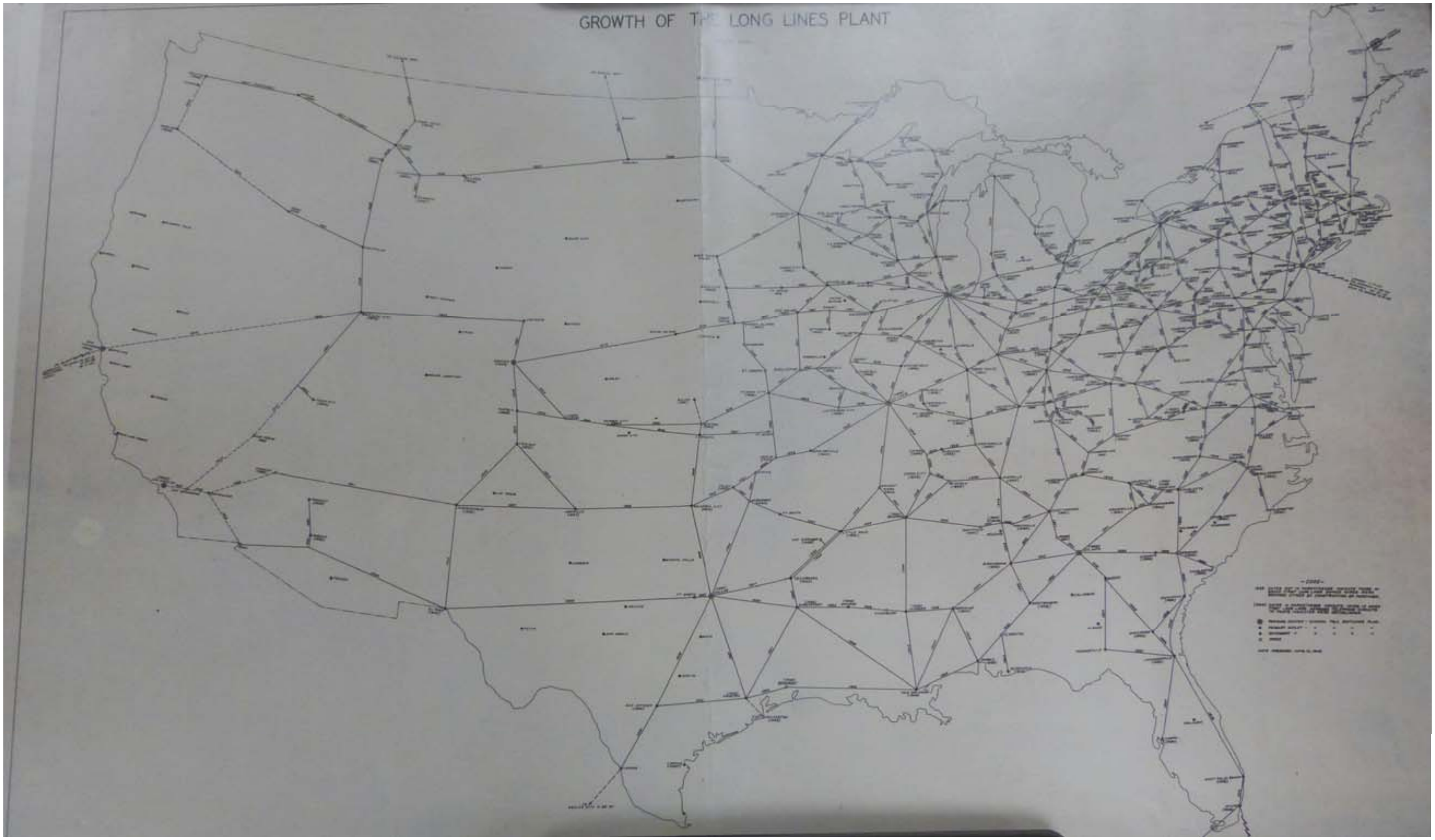

Note: Employment-weighted mean of thesaurus section task content based on verbs from time-invariant occupation descriptions from the 1991 DOTs.

# Figure A10



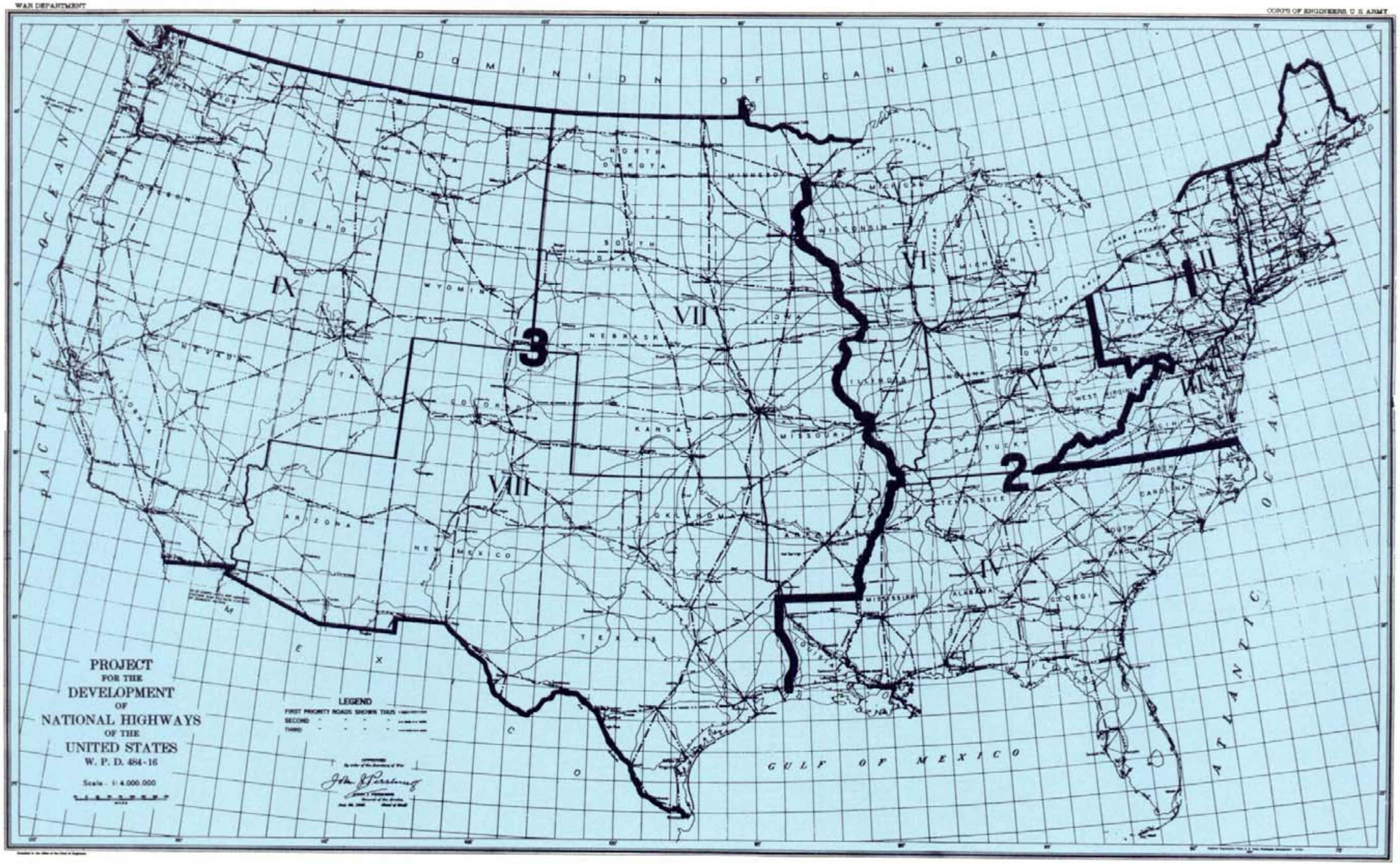

Note: Employment-weighted mean of thesaurus section task content based on verbs from time-invariant occupation descriptions from the 1991 DOTs.

## Map A1: American Telephone and Telegraph Company (AT&T) Long Distance Network




Source: American Telephone and Telegraph Company (AT&T), New Jersey.

## Map A2: Pershing Map 1922




Source: Department of Transportation (1976) *America's Highways 1776-1976: A History of the Federal Aid Program* , Washington, U. S. Government Printing Office