Accounting for Changes in Between-Group Inequality*

Ariel Burstein
UCLA

Eduardo Morales
Princeton University

Jonathan Vogel
Columbia University

December 31, 2013

PRELIMINARY AND INCOMPLETE

Abstract

We provide a framework with multiple worker types (e.g. gender, age, education), to decompose changes in aggregated and disaggregated between-group inequality into changes in (i) the supply of each worker type, (ii) the importance of different tasks, (iii) the extent of computerization, and (iv) other labor-specific productivities (a residual to match observed relative wages). The model features three forms of comparative advantage: between worker types and computers, worker types and tasks, and computers and tasks. We parameterize the model to match observed changes in worker type allocation and wages in the United States between 1984 and 2003. The combination of changes in the importance of tasks and computerization explain the majority of the rise in the skill premium as well as rising inequality across more disaggregated education types, whereas labor-specific productivity changes drive between-worker wage polarization.

*We thank Treb Allen, Costas Arkolakis, Davin Chor, Gene Grossman, Robert E. Lucas, Stephen Redding, Robert Shimer, and Nancy Stokey. Vogel thanks the Princeton International Economics Section for their support.

1 Introduction

The last few decades have witnessed pronounced changes in between-group wage inequality in the United States, both at the aggregate level—e.g. the rise in the college premium or the fall in the gender premium—and at the disaggregate level—e.g. the rise in wages for those groups of workers at the top and bottom of the wage distribution relative to those in the middle (i.e. between-group wage polarization). The same time period has been marked by dramatic changes in the economic environment, which a vast literature has argued have impacted inequality. These include rising relative supplies of educated workers and women in the workplace, phenomena such as structural transformation and international trade that changed the relative demand for workers across tasks, computerization, and other forms of labor-type-specific technical change; see e.g. Feenstra and Hanson (1999), Krusell et al. (2000), and Autor et al. (2003). In this paper, we provide a quantitative framework incorporating these forces and use it to decompose the sources of changes in aggregated and disaggregated between-group inequality between 1984-2003 in the United States.

Our framework features many types of workers (e.g. young men who dropped out of high school) and many types of capital equipment (e.g. computers) that are employed producing many tasks (e.g. health services), allowing us to study, respectively, disaggregated between-group inequality, the growth of computers relative to other forms of capital equipment, and the reallocation of workers across tasks. The productivity of a worker of a given type employed in a specific task and using a specific type of equipment has two components: a systematic component that is common to all workers of that type given that choice of task and equipment and an idiosyncratic component that is specific to that worker. The idiosyncratic component of productivity allows us to model workers' decisions as a tractable discrete choice problem, as in McFadden (1974), Eaton and Kortum (2002), and Hsieh et al. (2013). The systematic component of productivity—which varies with worker type, equipment type, and task—allows for three types of comparative advantage: between worker types and equipment types, between worker types and tasks, and between equipment types and tasks. Even though our framework imposes strong restrictions on micro-level production functions, at the aggregate level we obtain rich interactions between worker types, capital equipment types, and tasks, and we nest standard frameworks for studying between-group inequality.[1]

Comparative advantage (CA) shapes the allocation of labor to tasks and capital both

[1] In our discrete choice approach, factor allocation at a point in time provides information about aggregate elasticities between factors. An alternative approach, which is harder to implement in practice, is to specify a general aggregate production function with multiple cross elasticities between factors.

directly and indirectly. For instance, the fact that educated workers use computers relatively more than less educated workers can be generated by two distinct patterns of comparative advantage. First, if more educated workers have a CA with computers, they will use computers relatively more within tasks. Second, if more educated workers have a CA in the tasks in which computers have a CA, then they will disproportionately allocate to tasks in which all workers use computers more. In general, any aggregate pattern of factor allocation—workers to equipment, workers to tasks, or equipment to tasks—can be generated either directly (as in the first case) or indirectly (as in the second case) by comparative advantage. Given data on the allocation of workers to equipment types and tasks, we can identify the systematic component of productivity using the previous logic.

To quantify the sources of changes in between-group inequality, we allow for four types of aggregate changes over time in the economic environment: (i) the composition of the workforce across labor types ("labor supply"), which can be directly measured,[2] (ii) task preference parameters and productivity (which we refer to as "task shifters" because they generate shifts in employment across tasks), (iii) the productivity of using and producing different types of capital equipment ("equipment productivity"), and (iv) the productivity of each labor type ("labor productivity").

Through factor allocation, CA shapes the impact of these changes on relative wages. For example, a decline in the cost of producing computers, (iii), increases the relative wage of more educated workers if they have a CA using computers. On the other hand, by reducing the price of tasks in which comptuers have a CA, a decline in the cost of producing computers may reduce the relative wage of workers with CAs in tasks in which computers have a CA. Thus, our model is flexible enough so that computerization may increase the relative wage of workers who are relatively productive using computers and may reduce the relative wage of workers employed in tasks in which computers are particularly productive, as described by, e.g., Autor et al. (1998) and Autor et al. (2003).

We can infer task shifters and equipment productivity, (ii) and (iii), given data on changes over time in the allocation of workers to equipment types and tasks (e.g. the share of young men who dropped out of high school who use computers in health service tasks) without directly using data on wage changes; this avoids the critique of Acemoglu (2002) regarding previous work evaulating the role of capital-skill complementarity in the evolution of the skill premium using aggregate time series data. We infer labor productivity, (iv), as a residual—given (i), (ii), (iii)—to match data on relative average wages across worker types. Hence, given data and the structure of our model, we can decom-

[2] We treat education decisions as exogenous. See e.g. Restuccia and Vandenbroucke (2008) and Hsieh et al. (2013) for treatements of endogenous education.

pose changes in between-group inequality into our four components.

Implementing this methodology requires data over time on the allocation of worker types to equipment types and tasks. We obtain such data from the October Current Population Survey (CPS) Computer Use Supplement, which—in addition to a worker's characteristics (with which we group workers into 30 types based on age, gender, and education) and occupation (which we map to 20 tasks in the model)—provides information for certain years (1984, 1989, 1993, 1999, and 2003) on whether or not a worker has direct or hands on use of a computer—be it a personal computer, laptop, mini computer, or mainframe—at work. While this data allows us to estimate our model and conduct our decomposition, it is not without limitations, as we discuss in depth below. First, it imposes a narrow view of computerization. Second, it only provides information on the allocation of workers to one type of equipment, computers; we, therefore, must infer usage of the second type of equipment in the model, non-computer equipment.[3] Finally, we do not observe the share of each worker's time at work spent using computers.[4]

Our procedure uncovers some interesting patterns of comparative advantage. For example, because more educated workers use computers more intensively within occupations, we infer that more educated workers have a comparative advantage using computers. Whereas women use computers more intensively than men, this aggregate pattern is mostly driven by indirect comparative advantage: women are allocated to occupations in which all workers use computers more intensively. Hence, we infer that they have at most a weak comparative advantage using computers. Because all workers use computers intensively in occupations where thinking creatively and repetition are relatively important, we infer that computers have a comparative advantage in such occupations; similarly, computers have a comparative disadvantage in occupations where manual dexterity is relatively important. Finally, because educated workers are disproportionately employed in occupations where analyzing data is particularly important (given the type of capital used), we infer that they have a comparative advantage in such occupations; similarly, we infer that they have a comparative disadvantage in occupations in which repetition is particularly important. Our procedure also implies that computer productivity (the productivity of using and producing computers relative to non-computer equipment) rises rapidly over time because of the observed rise in computer usage, conditional on worker type and occupation. This finding is consistent with ample evidence showing a rapid de-

[3]To alleviate this concern, we show that the parameterized model is consistent with the allocation of equipment types across sectors in Appendix B, where we match tasks in the model with sectors in the data.

[4]To alleviate this concern, we show that our estimates are consistent with those obtained using an alternative data source which contains the share of hours worked using computers: The 2006 German *Qualification and Working Conditions* survey.

cline in the price of computers relative to all other equipment types and structures, which we do not directly use in our estimation procedure.[5]

At the aggregate level, we decompose changes in the skill premium and the gender wage gap. Over the full sample we find that the combination of changes in capital productivity (the rise in the productivity of computers) and task shifters (the expansion of occupations in which more educated workers have a comparative advantage) account for roughly 71% of the sum of the forces pushing the skill premium upward (the sum of task shifters, capital productivity, and labor productivity). For a value of the elasticity of substitution between tasks that is no larger than one, we find that the change in task shifters is the dominant force increasing the skill premium. Whereas the change in labor productivity is the only component of our decomposition that is estimated using changes in observed wages, it accounts for only roughly 29% of the sum of the forces pushing the skill premium upward. In other words, observable changes in the allocation of workers to tasks and computers explains the majority of the rise in the skill premium. On the other hand, the reduction in the gender gap is driven equally by changes in task shifters and labor productivity. Even though women are substantially more likely than men to use computers, changes in equipment productivity play almost no role in closing the gender gap because, as discussed above, we find that women have only a weak comparative advantage using computers. At the disaggregate level, we decompose changes in relative average wages across five education groups. Our results are consistent with those at the aggregate level: changes in labor productivity are not particularly important for explaining the rise in between education-group inequality. However, changes in labor productivity are important for other moments of the between-group earnings distribution: they have a *U-shaped* effect on relative wages (between-group wage polarization), decreasing wages of intermediate wage groups relative to the lowest and highest wage groups for our 15 gropus of men between 1989 and 2003.

We show that restricting the sources of comparative advantage—either by assuming away comparative advantage with tasks (for both workers and equipment) or comparative advantage with equipment (for both workers and tasks)—substantially alters the

[5]While the decline over time in the U.S. in the price of equipment relative to structures has been well-documented (see e.g. Greenwood et al. (1997)), we highlight that this is mostly driven by a decline in computer prices. For example, between 1984 and 2003: (i) the price of industrial equipment and transportation equipment relative to computers and peripheral equipment has risen by a factor of 32 and 34, respectively, (calculated using the BEA's Price Indexes for Private Fixed Investment in Equipment and Software by Type) and (ii) the quantity of computers and peripheral equipment relative to industrial equipment and transportation equipment rose by a factor of 35 and 33, respectively (calculated using the BEA's Chain-Type Quantity Indexes for Net Stock of Private Fixed Assets, Equipment and Software, and Structures by Type). We do not use equipment price or quantity data directly in our procedure in part because of quality-adjustment issues raised by, e.g. Gordon (1990).

results of our decomposition. This suggests that modeling all three sources of comparative advantage is important for decomposing changes in between-group inequality.

Finally, we extend our baseline closed economy model to incorporate international trade in capital equipment. We abstract from task trade because we do not have data on trade in occupational output; see e.g. Grossman and Rossi-Hansberg (2008) for a theoretical analysis of task trade and inequality and Feenstra and Hanson (1999) for an empirical treatment of offshoring and relative wages. We show analytically that, in the context of a gravity model of trade in final goods (consumption goods and each type of capital equipment) it is straightforward to solve for the impact of international trade on relative wages between any two years in our sample given domestic absorption shares and gravity elasticities, in addition to the parameter values recovered in our baseline exercise. Between 1984 and 2003, trade in capital equipment accounts for roughly 18% of the total impact of changes in capital productivity on the skill premium in the United States. Given our parameter estimates, the impact of equipment trade on relative wages would be substantially larger in countries that import a larger share of their computer equipment; see e.g. Burstein et al. (2013) and Parro (2013).[6]

We view our contributions as threefold. First, we nest four of the central mechanisms shaping relative wages proposed in the literature—labor supply, task shifters, capital-skill complementarity, and other labor specific productivity—and quantify their importance in the United States. Second, we identify the implications of changes in task shifters and capital productivity for relative wages using changes over time in the allocation of workers to equipment types and tasks rather than using data on changes in wages. Third, we analyze relative wages at high and low levels of worker aggregation. In doing so, our framewok extends Costinot and Vogel (2010) and Acemoglu and Autor (2011) to incorporate complementarity between worker types and other inputs, following Grossman et al. (2013). In using a model of this form to conduct quantitative exercises, our paper is closely related to Hsieh et al. (2013). Whereas they introduce wedges and focus on changes in the extent of misallocation over time, we study changes in relative wages.

Whereas we analyze the changing share of labor income allocated across labor types and of equipment income allocated across equipment types, Karabarbounis and Neiman (2013) analyze the changing aggregate share of capital and labor. We focus on between-group inequality in this paper. See e.g. Kambourov and Manovskii (2009), Huggett et al. (2011), Hornstein et al. (2011), and Helpman et al. (2012) for an analysis of within-group

[6]Our framework, which includes multiple types of capital equipment and comparative advantage between labor types and equipment types rationalizes the findings in Caselli and Wilson (2004), that countries with different distributions of education import different mixes of capital equipment.

inequality.

2 Environment

We consider an economy with n_Λ types of workers, indexed by $\lambda \in \{\lambda_1, ..., \lambda_{n_\Lambda}\}$, and n_K types of capital equipment, indexed by $\kappa \in \{\kappa_1, ..., \kappa_{n_K}\}$.[7] At time t, the endogenous stock of capital equipment κ is given by $K_t(\kappa)$ and the exogenous supply of labor λ is given by $L_t(\lambda)$. Individual workers are indexed by $\omega \in \Omega$, and the set of workers of type λ is given by $\Omega(\lambda) \subseteq \Omega$. Workers and capital equipment are employed by firms to produce tasks, indexed by $\sigma \in \{\sigma_1, ..., \sigma_{n_\Sigma}\}$.[8]

Tasks are combined to create a single final good according to a constant elasticity of substitution (CES) production function,

$$Y_t = \left(\sum_\sigma \mu_t(\sigma) \, y_t(\sigma)^{(\rho-1)/\rho} \right)^{\rho/(\rho-1)}, \tag{1}$$

where $\rho > 0$ is the elasticity of substitution across tasks, $y_t(\sigma) \geq 0$ is task σ output, and $\mu_t(\sigma) \geq 0$ is demand shifter for task σ. The final good is used for consumption, $C_t = Y_t(C)$, and capital equipment investment, $I_t(\kappa) = Y_t(\kappa)$. The resource constraint for the final good can be expressed as

$$Y_t = q_t(C) \, Y_t(C) + \sum_\kappa q_t(\kappa) \, Y_t(\kappa),$$

where $q_t(C)$ and $q_t(\kappa)$ denote the cost (in terms of the final good) of a unit of consumption and investment in equipment κ, respectively. The law of motion for capital κ is

$$K_{t+1}(\kappa) = (1 - dep(\kappa)) \, K_t(\kappa) + I_t(\kappa),$$

where $0 < dep(\kappa) < 1$ is the depreciation rate for equipment κ. Utility of the representative household is given by $\sum_{t=0}^{\infty} u_t(C_t)$.

The output of a worker $\omega \in \Omega(\lambda)$ employed in task σ and teamed with k units of capital κ is given by $[T_t(\lambda,\kappa,\sigma)\varepsilon_t(\omega,\kappa,\sigma)]^{1-\alpha(\sigma)} k^{\alpha(\sigma)}$.[9] The output elasticity of equipment in task σ is $\alpha(\sigma) \in (0,1)$. The systematic component of productivity affecting all workers

[7] In the quantitative implementation, we set $n_K = 2$ because of data restrictions.

[8] One could interpret task output in many different ways. In our empirical applications we take two approaches, matching tasks in the model with sectors or occupations in the data.

[9] One could interpret a type of capital equipment as any input other than labor. In the case of inputs that cannot be accumulated, the depreciation rate would be one.

of type λ when using equipment κ in task σ (henceforth "using κ in σ") at time t is

$$T_t(\lambda,\kappa,\sigma) \equiv T_{\lambda t}(\lambda)\, T_{\kappa t}(\kappa)\, T_{\sigma t}(\sigma)\, T(\lambda,\kappa,\sigma) \tag{2}$$

where $T_{xt}(x) \geq 0$ may vary over time for each $x = \lambda, \kappa, \sigma$—so that some worker types, capital types, and tasks may become more productive than others over time—whereas $T(\lambda,\kappa,\sigma) \geq 0$ is assumed constant across time and allows relative productivities to vary with factor allocation—so that, for example, some worker types may be relatively productive using some equipment types.[10] The idiosyncratic component of productivity that is specific to worker ω when using κ in σ is $\varepsilon_t(\omega,\kappa,\sigma)$. Each worker $\omega \in \Omega(\lambda)$ independently draws $\varepsilon_t(\omega,\kappa,\sigma)$ for each (κ,σ) pair from a Frechet distribution,

$$G(\varepsilon;\lambda) = \exp\left(-\varepsilon^{-\theta(\lambda)}\right),$$

where $\theta(\lambda) > 1$ is the shape parameter. A higher value of $\theta(\lambda)$ is associated with less dispersion.

All markets are perfectly competitive and all factors are freely mobile. The price of task σ and the rental rate of capital κ are given by $p_t(\sigma)$ and $r_t(\kappa)$, respectively.[11]

2.1 Alternative assumptions

Our framework imposes two strong restrictions on standard neoclassical task-level production functions. First, the marginal physical product of worker ω producing task σ is independent of the set of other workers employed in task σ. Second, a worker uses only one type of capital equipment. Here we show that the restriction that a worker uses only one type of equipment is not central to our theoretical analysis.

In the simplest possible extension, each worker is endowed with a number of units of time in each period t and a distinct vector of εs for each unit of time and chooses a (κ,σ)

[10] In our baseline model we assume that the time-varying components of productivity are separable between labor, tasks, and equipment types, to perform a clean decomposition between these forces. In ongoing work we allow for changes over time in the interaction between labor and task specific productivity (i.e. $T_{\lambda\sigma t}(\lambda,\sigma)$), or labor and equipment specific productivity (i.e. $T_{\lambda\kappa t}(\lambda,\kappa)$). In these cases we can only decompose changes in relative wages into three components. Our next draft will report the results under these extensions, which do not vary substantially from those in this draft. Moreover, the results are unchanged if we assume that the changing labor-specific component combines productivity, $T_{\lambda t}(\lambda)$, and a labor-specific distortion that creates a wedge between the wage received by the worker and the wage paid by the producer.

[11] Our analysis is unchanged if we introduce other homogeneous inputs that enter the production function multiplicatively, in which case our production function for worker output in task σ would correspond to value added after the optimal choice of other inputs.

for each unit of his time. In this case, all of our aggregate conditions would hold exactly; moreover, each worker would allocate his time across (potentially) many (κ, σ) pairs. Alternatively, consider an environment in which the output of worker $\omega \in \Omega(\lambda)$ in σ using $k(\omega, \kappa, \sigma)$ units of capital for each type of capital κ is $[T_t(\lambda, \sigma)\, \varepsilon_t(\omega, \sigma)]^{1-\alpha(\sigma)} K_t(\omega, \sigma)^{\alpha(\sigma)}$, where $K_t(\omega, \sigma) = \left[\sum_\kappa T_t(\lambda, \kappa, \sigma)\, k_t(\omega, \kappa, \sigma)^{\frac{\theta(\lambda)-1}{\theta(\lambda)}}\right]^{\frac{\theta(\lambda)}{\theta(\lambda)-1}}$, $T_t(\lambda, \sigma) \geq 0$ is the systematic productivity of any $\omega \in \Omega(\lambda)$ working in task σ, and $\varepsilon_t(\omega, \sigma)$ is distributed Frechet with shape parameter $\theta(\lambda)$. Under this alternative set of assumptions, we obtain a characterization of aggregate equilibrium outcomes as a function of the (endogenous) allocation of labor across (κ, σ) that is identical to the present model. However, each worker uses (potentially) many types of equipment.

2.2 Relation to alternative frameworks

Whereas our framework imposes strong restrictions on production functions at the level of $(\lambda, \kappa, \sigma)$, at the aggregate level we obtain rich interactions between worker types, capital equipment types, and tasks, in the sense that our model nests two standard frameworks for studying between-group inequality.

The first of these, which is refered to as *the canonical model,* is the "central organizing framework of the voluminous recent literature studying changes in the returns to skills and the evolution of earnings inequality" (Acemoglu and Autor, 2011). In this framework, the aggregate production function is given by

$$Y_t = \left[A_{1t} L_t(\lambda_1)^{\frac{\rho-1}{\rho}} + A_{2t} L_t(\lambda_2)^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}}$$

where A_{1t} and A_{2t} are parameters. We obtain the above aggregate production function—and, therefore, the same relative wage—under a number of restrictions on our framework: (i) there are exactly two labor types, (ii) the capital equipment share of production is zero in each task, and (iii) labor type λ_i is only productive in task i.

The second of these is an extension of the canonical model that incorporates capital-skill complementarity—see e.g. Krusell et al. (2000)—where the aggregate production function, under the restriction that the elasticity of the CES nest of λ_1 and κ_1 is one, is given by

$$Y_t = \left[A_{1t} \left(L_t(\lambda_1)^{1-\alpha(\sigma_2)} K_t(\kappa_1)^{\alpha(\sigma_2)}\right)^{\frac{\rho-1}{\rho}} + A_{2t} L_t(\lambda_2)^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}}$$

where A_{1t} and A_{2t} are parameters. We obtain the above aggregate production function under a number of restrictions on our framework: (i) there are exactly two labor types, (ii) there is one type of capital equipment, (iii) the capital equipment share of production is zero in task two, and (iii) and labor type λ_i is only productive in task i. Hence, while the restrictions we impose facilitate our analysis in high-dimensional environments, they do not limit our framework in the low-dimensional analyses considered previously.

3 Equilibrium and mechanisms

In this section we characterize the equilibrium of the model, first in partial equilibrium—given $p_t(\sigma)$ and $r_t(\kappa)$—in Section 3.1 and then in general equilibrium in Section 3.2. We then provide the system of equations with which to calculate changes in wages along a balanced growth path in Section 3.3. Finally, we discuss how comparative advantage shapes factor allocation and wage changes in Section 3.4.

3.1 Partial equilibrium

The producer's zero profit condition implies that a worker $\omega \in \Omega(\lambda)$ teamed with k units of capital κ and producing task σ earns a wage given by $p_t(\sigma)\, k^{\alpha(\sigma)} \left[T_t(\lambda,\kappa,\sigma)\varepsilon_t(\omega,\kappa,\sigma)\right]^{1-\alpha(\sigma)} - kr_t(\kappa)$. If a worker chooses to work with κ in σ, he chooses the amount of κ to maximize his wage. Given the optimal capital decision, the worker's wage is simply $\tau_t(\lambda,\kappa,\sigma)^{1/\theta(\lambda)}\, \varepsilon_t(\omega,\kappa,\sigma)$, where

$$\tau_t(\lambda,\kappa,\sigma) \equiv \left[T_t(\lambda,\kappa,\sigma)\left(1-\alpha(\sigma)\right)\left(\frac{\alpha(\sigma)}{r_t(\kappa)}\right)^{\frac{\alpha(\sigma)}{1-\alpha(\sigma)}} p_t(\sigma)^{\frac{1}{1-\alpha(\sigma)}}\right]^{\theta(\lambda)}. \tag{3}$$

With perfectly competitive labor markets, worker ω is employed in the task and teamed with the type of capital that maximizes his wage. Given the distributional assumption on idiosyncratic productivity, the probability that a randomly sampled worker, $\omega \in \Omega(\lambda)$, uses κ in σ is simply

$$\pi_t(\lambda,\kappa,\sigma) = \frac{\tau_t(\lambda,\kappa,\sigma)}{\sum_{\kappa',\sigma'} \tau_t(\lambda,\kappa',\sigma')}. \tag{4}$$

The distributional assumption also implies that the average wage of workers $\omega \in \Omega(\lambda)$ teamed with κ in σ does not vary across κ,σ and is given by

$$w_t(\lambda) = \gamma(\lambda)\left(\sum_{\sigma,\kappa} \tau_t(\lambda,\kappa,\sigma)\right)^{1/\theta(\lambda)}, \tag{5}$$

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

3.2 General equilibrium

In any period, task markets must clear,

$$\mu_t\left(\sigma\right)\left(\frac{p_t\left(\sigma\right)}{P_t}\right)^{1-\rho} E_t = \frac{1}{1-\alpha\left(\sigma\right)}\zeta_t\left(\sigma\right), \tag{6}$$

where $\zeta_t\left(\sigma\right)$ denotes total labor income in task σ in period t, $\zeta\left(\sigma\right) = \sum_{\lambda,\kappa} w_t\left(\lambda\right) L_t\left(\lambda\right) \pi_t\left(\lambda,\kappa,\sigma\right)$, so that the right-hand side condition (6) is total income in task σ. The left-hand side is expenditure on task σ, where E_t is total income and $P_t = \left(\sum_\sigma \mu_t\left(\sigma\right) p_t\left(\sigma\right)^{1-\rho}\right)^{1/(1-\rho)}$ is the price of the final good.

The prices of consumption and each type of capital equipment are simply

$$P_t\left(C\right)/q_t\left(C\right) = P_t\left(\kappa\right)/q_t\left(\kappa\right) = P_t. \tag{7}$$

The dynamic Euler equation for the accumulation of capital κ and equation (7) give

$$u'_{C,t}\frac{q_t\left(C\right)}{q_t\left(\kappa\right)} = u'_{C,t+1}\frac{q_{t+1}\left(C\right)}{q_{t+1}\left(\kappa\right)}\left(\frac{r_{t+1}\left(\kappa\right)}{P_{t+1}\left(\kappa\right)} + \left(1-dep\left(\kappa\right)\right)\right), \tag{8}$$

where $u'_{C,t}$ is the marginal utility of consumption at time t.

3.3 Wage changes

In this section we provide the system of equations with which to calculate wage changes along a balanced growth path (BGP). Between any two time periods, t_0 and t_1, we use equation (5) to express changes in wages as

$$\hat{w}(\lambda) = \hat{T}_\lambda(\lambda)\left\{\sum_{\kappa,\sigma}\left[\left(\hat{r}(\kappa)^{-\alpha(\sigma)}\hat{p}(\sigma)\right)^{\frac{1}{1-\alpha(\sigma)}}\hat{T}_\kappa(\kappa)\hat{T}_\sigma(\sigma)\right]^{\theta(\lambda)}\pi_t(\lambda,\kappa,\sigma)\right\}^{1/\theta(\lambda)} \tag{9}$$

where $\hat{x} \equiv x_{t_1}/x_{t_0}$. Changes in task prices between t_0 and t_1 are determined by the task market clearing conditions, equation (6),

$$\frac{\hat{\mu}\left(\sigma\right)}{\hat{\mu}\left(\sigma_1\right)}\left(\frac{\hat{p}\left(\sigma\right)}{\hat{p}\left(\sigma_1\right)}\right)^{1-\rho} = \frac{\hat{\zeta}\left(\sigma\right)}{\hat{\zeta}\left(\sigma_1\right)} \tag{10}$$

where

$$\hat{\zeta}(\sigma) = \frac{\sum_{\lambda,\kappa} w_{t_0}(\lambda) L_{t_0}(\lambda) \pi_{t_0}(\lambda,\kappa,\sigma) \hat{w}_t(\lambda) \hat{L}_t(\lambda) \hat{\pi}_t(\lambda,\kappa,\sigma)}{\sum_{\lambda,\kappa} w_{t_0}(\lambda) L_{t_0}(\lambda) \pi_{t_0}(\lambda,\kappa,\sigma)} \tag{11}$$

and where, using equation (4), changes in allocations are given by

$$\hat{\pi}(\lambda,\kappa,\sigma) = \frac{\left(\hat{T}_\lambda(\lambda)\hat{T}_\kappa(\kappa)\hat{T}_\sigma(\sigma)\left(\hat{r}(\kappa)^{-\alpha(\sigma)}\hat{p}(\sigma)\right)^{\frac{1}{1-\alpha(\sigma)}}\right)^{\theta(\lambda)}}{\sum_{\sigma',\kappa'}\left(\hat{T}_\lambda(\lambda)\hat{T}_\kappa(\kappa')\hat{T}_\sigma(\sigma')\left(\hat{r}(\kappa')^{-\alpha(\sigma)}\hat{p}(\sigma')\right)^{\frac{1}{1-\alpha(\sigma)}}\right)^{\theta(\lambda)}\pi_t(\lambda,\kappa',\sigma')}. \tag{12}$$

Finally, we must determine changes in rental rates. We assume that in periods t_0 and t_1 the economy is in a BGP in which the real interest rate, $\frac{u'_{C,t+1}}{u'_{C,t}}$, and the growth rate of relative productivity, $\frac{q_{t+1}(C)}{q_t(C)}\Big/\frac{q_{t+1}(\kappa)}{q_t(\kappa)}$, are constant over time. Conditions (7) and (8) provide the following relationship between changes in rental rates and prices across two BGPs,

$$\hat{r}(\kappa) = \hat{P}(\kappa) = \hat{q}(\kappa)\hat{P}(C)/\hat{q}(C),$$

so that changes in relative rental rates are determined solely by changes in production costs,

$$\frac{\hat{r}(\kappa')}{\hat{r}(\kappa)} = \frac{\hat{q}(\kappa')}{\hat{q}(\kappa)}. \tag{13}$$

This leads to the following proposition.

Proposition 1. *Given values of $w_t(\lambda) L_t(\lambda)$, $\pi_t(\lambda,\kappa,\sigma)$, ρ, $\alpha(\sigma)$, and $\theta(\lambda)$ for all λ,κ,σ, equations (9)-(13) determine changes in wages for any changes in technology, $\hat{q}(\kappa)$, $\hat{\mu}(\sigma)$, and $\hat{T}_x(x)$ for all $x = \lambda,\kappa,\sigma$, or factor supply, $\hat{L}(\lambda)$, between two BGPs.*

3.4 Mechanisms

Comparative advantage. There are three types of comparative advantage in our model: *(i)* between labor and equipment, *(ii)* between equipment and tasks, and *(iii)* between labor and tasks. We define comparative advantage between labor and equipment as follows: λ' has a comparative advantage (relative to λ) using equipment κ' (relative to κ) in σ if $\frac{T(\lambda',\kappa',\sigma)}{T(\lambda',\kappa,\sigma)} \geq \frac{T(\lambda,\kappa',\sigma)}{T(\lambda,\kappa,\sigma)}$. We define comparative advantage between labor and task and between equipment and task similarly. Comparative advantage has strong implications for factor allocation. For instance, note that if λ' has a comparative advantage (relative to λ) using κ' (relative to κ) in σ, then condition (4) implies $\frac{\pi(\lambda',\kappa',\sigma)}{\pi(\lambda',\kappa,\sigma)} \geq \frac{\pi(\lambda,\kappa',\sigma)}{\pi(\lambda,\kappa,\sigma)}$.

To better understand the role of comparative advantage on factor allocation, consider the following feature of the data we describe in Section 5.1: the share of workers using computers (κ') relative to other non-computer equipment (κ) is higher for college educated workers (λ') than for high school educated workers (λ). Two distinct patterns of comparative advantage could generate this feature of the data. First, college educated workers could have a comparative advantage using computers within tasks, i.e. $\frac{T(\lambda',\kappa',\sigma)}{T(\lambda',\kappa,\sigma)} \geq \frac{T(\lambda,\kappa',\sigma)}{T(\lambda,\kappa,\sigma)}$ for all σ. Alternatively, they could have a comparative advantage in the tasks (σ') in which computers have a comparative advantage, e.g. $\frac{T(\lambda',\kappa_0,\sigma')}{T(\lambda',\kappa_0,\sigma)} \geq \frac{T(\lambda,\kappa_0,\sigma')}{T(\lambda,\kappa_0,\sigma)}$ for all κ_0 and $\frac{T(\lambda_0,\kappa',\sigma')}{T(\lambda_0,\kappa,\sigma')} \geq \frac{T(\lambda_0,\kappa',\sigma)}{T(\lambda_0,\kappa,\sigma)}$ for all λ_0. These explanations can be separated in the data as follows. In the first case, college educated workers would use computers relatively more than high school educated workers *within tasks*. In the second case, employment composition *across tasks* would be key: college educated workers would be employed relatively more in tasks in which computers are used more frequently by all workers. In general, any aggregate pattern of factor allocation—workers to equipment, workers to tasks, or equipment to tasks—can be generated either directly (as in the first case) or indirectly (as in the second case) by comparative advantage, and disaggregated data on allocations would allow us to separate these two distinct explanations.

Comparative advantage and wage changes. According to the wage equation (9), changes in wages depend linearly—for given prices—on changes in worker-specific productivities, $T_{\lambda t}(\lambda)$, and are a CES combination of changes in prices, $r(\kappa)$ and $p(\sigma)$, the productivity of using equipment types, $T_{\kappa t}(\kappa)$, and the productivity of employment in tasks, $T_{\sigma t}(\sigma)$, where the weight given to changes in each of these components depends, through factor allocation $\pi_t(\lambda, \kappa, \sigma)$, on comparative advantage. Hence, each of the three types of comparative advantage present in our model plays a central role in shaping the impact of changes in the economic environment on relative wages.

Consider, for example, the impact on relative wages of a reduction in the cost of producing computers, $q(\kappa)$. There are two forces that shape the response of relative wages. The first is driven by direct comparative advantage between workers and computers. If λ workers have a comparative advantage using computers, a reduction in $q(\kappa)$—which reduces $r(\kappa)$—increases their relative wages. Intuitively, a fall in computer prices helps workers who tend to use computers relatively more within tasks.

The second force shaping the response of relative wages is driven by the indirect comparative advantage between workers and computers (i.e. worker-task, and equipment-task comparative advantage). If λ workers have a comparative advantage in tasks in which computers have a comparative advantage, then the impact of a reduction in $q(\kappa)$—

which also reduces $p(\sigma)$ in the tasks in which computers have a comparative advantage—on relative wages depends on the value of ρ. If $\rho < 1$, then the decline in task prices is greater than the rise in $r(\kappa)^{-\alpha(\sigma)}$, and the relative wage of λ workers falls (see expression 9); whereas if $\rho > 1$, then the decline in task prices is less than the rise in $r(\kappa)^{-\alpha(\sigma)}$, and the relative wage of λ workers rises. Intuitively, if $\rho < 1$, then a fall in the price of computers decreases employment (and hence hurts workers) in the tasks in which computers have a comparative advantage. If $\rho > 1$ then a fall in the price of computers increases employment (and hence benefits workers) in the tasks in which computers have a comparative advantage. Similar intuition applies to the impact on relative wages of changes in other primitives, which we consider in our decomposition.

These forces capture some of the standard mechanisms shaping relative wages described in the literature. Summarizing this literature, Autor et al. (1998) discuss two channels through which computers may influence relative labor demand: (i) computers may directly substitute for human judgment and labor and (ii) computers may increase the returns to the creative use of greater available information. Specifically, Autor et al. (2003) find that *"(i)* computer capital substitutes for workers in performing cognitive and manual tasks that can be accomplished by following explicit rules" and *"(ii)* complements workers in performing nonroutine problem solving and complex communications tasks." By incorporating all three types of comparative advantage, our model is theoretically consistent with these findings and, therefore, has the potential to account for the fact that the fall in the price of computers can help some worker types who are employed in tasks in which computers are prevalent while hurting others.[12] Moreover, when $\rho < 1$ our empirical results are consistent with the findings in Autor et al. (2003).[13]

4 Decomposing changes in between group inequality

We aim to perform a decomposition quantifying the direct contribution for changes in relative wages between time periods t_0 and t_1 of

i. changes in *labor supply*, $\hat{L}(\lambda)$,

[12] In the model presented in Acemoglu and Autor (2011), capital equipment only influences relative demand through the indirect comparative advantage channel because their model abstracts from worker-capital comparative advantage. Hence, an increase in the computer stock must hurt worker types that are disproportionately employed in tasks in which computers are prevalent.

[13] Specifically, we find that computers have a comparative advantage in tasks in which repetition is particularly important. Hence, when $\rho < 1$ a fall in computer prices hurts (relatively) workers with a comparative advantage in such tasks. Moreover, we find that more educated workers (who tend to peform tasks in which thinking creatively is important) have a comparative advantage using computers. Hence, educated workers benefit (relatively) from a fall in computer prices.

ii. changes in *labor productivity*, $\hat{T}_\lambda(\lambda)$,

iii. changes in *capital productivity*, $\hat{T}_\kappa(\kappa)$ and $\hat{q}(\kappa)$, and

iv. changes in *task shifters*, $\hat{T}_\sigma(\sigma)$ and $\hat{\mu}(\sigma)$.

Specifically, the direct contribution of changes in labor supply (for example) could be measured by solving for movements in the log of relative wages that result from changing labor supplies from $L_{t_0}(\lambda)$ to $L_{t_1}(\lambda)$, holding all other parameters fixed at their t_0 levels. We could similarly determine the direct contributions of changes in labor productivity, capital productivity, and task shifters. Of course, these direct contributions need not sum up to the change in log relative wages that results from changing all parameters from their t_0 to their t_1 levels because of interaction effects; see e.g. Rothe (2012a,b). In practice, however, these interaction effects are very small in Section 6.

According to Proposition 1, we can conduct this decomposition exercise if we have values of parameters ρ, $\alpha(\sigma)$, and $\theta(\sigma)$, values of labor payments and allocations in period t, and measures of $\hat{L}(\lambda)$, $\hat{T}_\lambda(\lambda)$, $\hat{T}_\sigma(\sigma)$, $\hat{\mu}(\sigma)$, $\hat{T}_\kappa(\kappa)$, and $\hat{q}(\kappa)$. In the remainder of this paper, we impose a common equipment intensity in each task, $\alpha(\sigma) = \alpha$ for all σ,[14] and a common dispersion of idiosyncratic productivities across worker types, $\theta(\lambda) = \theta$ for all λ.

Given the data that we use, described in section 5.1, we are unable to obtain estimates of $\hat{T}_\sigma(\sigma)$, $\hat{\mu}(\sigma)$, $\hat{T}_\kappa(\kappa)$, and $\hat{q}(\kappa)$. Nevertheless, we now show that we can perform our decomposition using transformed variables that we are able to estimate as described in Section 5. These transformed variables are defined as follows. Combining equations (2) and (3), we have

$$\tau_t(\lambda,\kappa,\sigma) = \tau_{\lambda t}(\lambda)\,\tau_{\kappa t}(\kappa)\,\tau_{\sigma t}(\sigma)\,\tau(\lambda,\kappa,\sigma) \tag{14}$$

where

$$\begin{aligned}
\tau_{\lambda t}(\lambda) &\equiv \left[\frac{T_{\lambda t}(\lambda)}{\prod_{t'=1}^{T} T_{\lambda t'}(\lambda)}\right]^{\theta} \\
\tau_{\kappa t}(\kappa) &\equiv \left[\frac{T_{\kappa t}(\kappa)\, r_t(\kappa)^{\frac{-\alpha}{1-\alpha}}}{\prod_{t'=1}^{T} T_{\kappa t'}(\kappa)\, r_{t'}(\kappa)^{\frac{-\alpha}{1-\alpha}}}\right]^{\theta} \\
\tau_{\sigma t}(\sigma) &\equiv \left[\frac{T_{\sigma t}(\sigma)\, p_t(\sigma)^{\frac{1}{1-\alpha}}}{\prod_{t'=1}^{T} T_{\sigma t'}(\sigma)\, p_{t'}(\sigma)^{\frac{1}{1-\alpha}}}\right]^{\theta}
\end{aligned}$$

[14] We aim relax this assumption in future drafts.

and

$$\tau\left(\lambda,\kappa,\sigma\right)\equiv\left[\alpha^{\frac{\alpha}{1-\alpha}}\left(1-\alpha\right)T(\lambda,\kappa,\sigma)\prod_{t'=1}^{T}\left(T_{\lambda t'}\left(\lambda\right)T_{\kappa t'}\left(\kappa\right)r_{t'}\left(\kappa\right)^{\frac{-\alpha}{1-\alpha}}T_{\sigma t'}\left(\sigma\right)p_{t'}\left(\sigma\right)^{\frac{1}{1-\alpha}}\right)\right]^{\theta}.$$

We now describe how to conduct the decomposition described above for given values of (i) ρ,θ, and α; (ii) $L_t\left(\lambda\right)/L_t\left(\lambda_1\right)$; (iii) $\psi\tau\left(\lambda,\kappa,\sigma\right)$ for an arbitrary constant ψ; and (iv) $\tau_{\lambda t}\left(\lambda\right)/\tau_{\lambda t}\left(\lambda_1\right)$, $\tau_{\sigma t}\left(\sigma\right)/\tau_{\sigma t}\left(\sigma_1\right)$, and $\tau_{\kappa t}\left(\kappa\right)/\tau_{\kappa t}\left(\kappa_1\right)$ for all λ,κ,σ and $t=t_0,t_1$. Given $(i)-(iv)$, we construct $\pi_{t_0}\left(\lambda,\kappa,\sigma\right)$ and $w_{t_0}\left(\lambda\right)L_{t_0}\left(\lambda\right)/w_{t_0}\left(\lambda_1\right)L_{t_0}\left(\lambda_1\right)$ using equations (4), (5), and (14). Given (ii) we construct $\hat{L}\left(\lambda\right)/\hat{L}\left(\lambda_1\right)$ and given (iv) we construct $\hat{\tau}_{\lambda}\left(\lambda\right)/\hat{\tau}_{\lambda}\left(\lambda_1\right)$, $\hat{\tau}_{\kappa}\left(\kappa\right)/\hat{\tau}_{\kappa}\left(\kappa_1\right)$, and $\hat{\tau}_{\sigma}\left(\sigma\right)/\hat{\tau}_{\sigma}\left(\sigma_1\right)$. Using these, we conduct each exercise as follows. Details are provided in the Appendix.

Labor supply. To quantify the direct impact of changes in labor supply, we set $\hat{T}_{\lambda}\left(\lambda\right)$, $\hat{T}_{\kappa}\left(\kappa\right)$, $\hat{T}_{\sigma}\left(\sigma\right)$, $\hat{\mu}\left(\sigma\right)$, and $\hat{q}\left(\kappa\right)$ all equal to one for each $\left(\lambda,\kappa,\sigma\right)$, impose $\frac{\hat{L}(\lambda)}{\hat{L}(\lambda_1)}=\frac{L_{t_1}(\lambda)/L_{t_0}(\lambda)}{L_{t_1}(\lambda_1)/L_{t_0}(\lambda_1)}$, and use equations (9)-(13) to solve for changes in relative wages, as summarized by Lemma 1 in the Appendix.

Labor productivity. To quantify the direct impact of changes in labor productivity, we set $\hat{L}\left(\lambda\right)$, $\hat{T}_{\kappa}\left(\kappa\right)$, $\hat{T}_{\sigma}\left(\sigma\right)$, $\hat{\mu}\left(\sigma\right)$, and $\hat{q}\left(\kappa\right)$ all equal to one for each $\left(\lambda,\kappa,\sigma\right)$, impose $\hat{T}_{\lambda}\left(\lambda\right)/\hat{T}_{\lambda}\left(\lambda_1\right)=\left(\hat{\tau}_{\lambda}\left(\lambda\right)/\hat{\tau}_{\lambda}\left(\lambda_1\right)\right)^{1/\theta}$, and use equations (9)-(13) to solve for changes in relative wages, as summarized by Lemma 2 in the Appendix.

Capital productivity. While both $\hat{T}_{\kappa}\left(\kappa\right)$ and $\hat{q}\left(\kappa\right)$ shape capital productivity, we cannot separately recover $\hat{T}_{\kappa}\left(\kappa\right)$ and $\hat{q}\left(\kappa\right)$ from $\hat{\tau}_{\kappa}\left(\kappa\right)$ without additional data. However, changes in relative wages depend on $\hat{T}_{\kappa}\left(\kappa\right)$ and $\hat{q}\left(\kappa\right)$ only through $\frac{\hat{\tau}_{\kappa}(\kappa')}{\hat{\tau}_{\kappa}(\kappa)}=\left(\frac{\hat{T}_{\kappa}(\kappa')}{\hat{T}_{\kappa}(\kappa)}\left(\frac{\hat{q}(\kappa')}{\hat{q}(\kappa)}\right)^{\frac{-\alpha}{1-\alpha}}\right)^{\theta}$. Hence, to quantify the direct impact of changes in labor productivity, we set $\hat{L}\left(\lambda\right)$, $\hat{T}_{\lambda}\left(\lambda\right)$, $\hat{T}_{\sigma}\left(\sigma\right)$, and $\hat{\mu}\left(\sigma\right)$ all equal to one for each $\left(\lambda,\kappa,\sigma\right)$, impose $\frac{\hat{T}_{\kappa}(\kappa')}{\hat{T}_{\kappa}(\kappa)}\left(\frac{\hat{q}(\kappa')}{\hat{q}(\kappa)}\right)^{-\alpha/(1-\alpha)}=\left(\frac{\hat{\tau}_{\kappa}(\kappa')}{\hat{\tau}_{\kappa}(\kappa)}\right)^{1/\theta}$, and use equations (9)-(13) to solve for changes in relative wages, as summarized by Lemma 2 in the Appendix. As long as we are not interested in separating the effects of changes in the costs of producing equipment, $\hat{q}\left(\kappa\right)$, from changes in the productivity using equipment, $\hat{T}_{\kappa}\left(\kappa\right)$, our estimates of $\hat{\tau}_{\kappa}\left(\kappa\right)$ are sufficient to conduct this exercise.

Task shifters. As in the capital productivity component, we cannot separate $\hat{T}_{\sigma}\left(\sigma\right)$ and $\hat{\mu}\left(\sigma\right)$ from $\hat{\tau}_{\sigma}\left(\sigma\right)$ without additional data. Nevertheless, we can quantify the direct impact of changes in task shifters by setting $\hat{L}\left(\lambda\right)$, $\hat{T}_{\lambda}\left(\lambda\right)$, $\hat{T}_{\kappa}\left(\kappa\right)$, and $\hat{q}\left(\kappa\right)$ all equal to one for each $\left(\lambda,\kappa,\sigma\right)$, and using equations (9)-(13) to solve for changes in relative wages for given $\hat{\tau}_{\sigma}\left(\sigma\right)$ and for given changes in incomes accruing in each task, $\hat{\zeta}\left(\sigma\right)$, as summarized by Lemma 4 in the Appendix. Whereas changes in the relative importance of tasks in the production of the unique final good, $\hat{\mu}\left(\sigma\right)$, and changes in the productivity of tasks,

$\hat{T}_\sigma(\sigma)$, have different implications for changes in task prices, their combined effects on relative wages is summarized in $\hat{\tau}_\sigma(\sigma)$ and $\hat{\zeta}(\sigma)$. Hence, estimates of $\hat{\tau}_\sigma(\sigma)$ and $\hat{\zeta}(\sigma)$ are sufficient to quantify the direct impact of changes in task shifters.

We summarize our previous discussion in the following proposition.

Proposition 2. *Given ρ, θ, α, and—for all λ, κ, σ and $t = t_0, t_1$—$L_t(\lambda) / L_t(\lambda_1)$, $\psi\tau(\lambda, \kappa, \sigma)$ for an arbitrary constant $\psi \neq 0$, and $\tau_{xt}(x) / \tau_{xt}(x_1)$ for all $x = \lambda, \kappa, \sigma$, equations (9)-(13) provide an algorithm to solve for the direct contributions, between t_0 and t_1, of (i) $\hat{L}(\lambda)$, (ii) $\hat{T}_\lambda(\lambda)$, (iii) $\hat{T}_\sigma(\sigma)$ and $\hat{\mu}(\sigma)$, and (iv) $\hat{T}_\kappa(\kappa)$ and $\hat{q}(\kappa)$.*

5 Connecting model and data

In what follows, we describe how we map our model to data in order to perform the decomposition described in Section 4. We first outline our main data sources. Next, we describe our model parameterization. Some parameters are assigned and some are estimated. We then summarize two features of our parameterization that are a key input for our decomposition: comparative advantage and changes in task shifters and labor and capital productivities.

5.1 Data

Our primary data sources are the October CPS Supplement (October Supplement) and the March Current Population Survey (March CPS). We restrict our sample to those workers at least 16 years old reporting positive (paid) hours worked. Here we briefy describe our use of these sources; we provide further details in Appendix A.2.

March CPS. The March CPS provides measures of the prior year's annual earnings, weeks worked, and hours worked per week over our timeframe. We use the March CPS to form a sample—for each worker type—of hours worked and income. Our measure of labor supply, $L_t(\lambda)$, is hours worked within labor type λ, and our measure of the average (hourly) wage, $w_t(\lambda)$, is total income (including the sum of labor, business, and farm income in the previous year) divided by total hours worked within λ; see Appendix A.2 for our treatment of top-coded income. In addition to the two genders, we group workers into three age categories—16-30, 31-43, and 44 and older—and five education categories—high school dropouts (HSD), high school graduates (HSG), some college (SMC), completed college (CLG), and graduate training (GTC)—yielding a total of thirty

		1984	1989	1993	1997	2003
All		25.1	36.6	46.3	50.5	57.0
Gender	Men	21.6	31.3	40.6	44.5	51.6
	Women	30.6	44.2	54.4	59.0	64.3
Age	16-30	25.4	35.0	42.2	45.8	48.8
	31-43	29.1	41.7	50.9	54.2	59.8
	44 +	20.3	32.1	44.6	50.2	59.5
Education	HSD	5.2	7.3	10.4	12.2	15.9
	HSG	19.5	28.2	33.9	36.2	40.6
	SMC	32.5	45.8	53.3	56.6	59.5
	CLG	42.6	58.3	70.3	75.8	83.5
	GTC	42.6	58.4	71.3	79.1	87.8

Table 1: The probability of using a computer, weighted by hours worked

HSD: high school dropout; HSG: high school graduate; SMC: some college; CLG: college; GTC: graduate training

labor types.[15] Because we use questions in the March CPS that refer to the previous year, year t's March CPS refers to year $t-1$. To create labor supply and average wages for year t, we average the March CPS for years t, $t+1$, and $t+2$ to reduce measurement error.

October Supplement. In our baseline exercises, we map tasks in the model to occupations in the data. In the Appendix we conduct our decomposition mapping tasks in the model to sectors in the data. We aggregate up to twenty occupations.[16] See Table 10 in Appendix A.2 for a list of these occupations. In 1984, 1989, 1993, 1997, and 2003, the October Supplement asked respondents whether they "have direct or hands on use of computers at work," "directly use a computer at work," or "use a computer at/for his/her/your main job." Using a computer at work refers only to "direct" or "hands on" use of a computer with typewriter like keyboards, whether a personal computer, laptop, mini computer, or mainframe.

Table 1 summarizes the fraction of workers using a computer at work for several categories of workers (weighted by hours) for each year of the October Supplement. There are a few things to note. First, the share of workers using computers rises over time. Second, the share of women using computers is higher than the share of men in each year. Finally, across every education category and in every year, more educated workers

[15]Our quantitative results on the sources of changes in aggregate measures of between-group inequality are largely robust to further aggregating labor types. For instance, we obtain similar results for the skill premium (or the gender premium) with only two labor types: college educated and non-college educated workers (or men and women). Of course, at this level of aggregation we cannot speak to changes in more disaggregated measures of between-group inequality.

[16]While we could map tasks in the model to (occupation, sector) pairs, in practice the data would become sparse (unless we reduced the number of occupations or sectors), in the sense that there would be many $(\lambda, \kappa, \sigma)$ and t for which $\pi_t(\lambda, \kappa, \sigma) = 0$.

are more likely to use computers than less educated workers, and the gap is substantial across all comparisons except between those with college degrees and graduate training. All of these results are robust to conditioning on occupation of employment with one exception. The gap between genders becomes significantly smaller when we control for occupation; in practice, women are employed disproportionately in occupations in which computers are used more intensively by all workers.

Given that the October Supplement only provides information on the use of one form of capital equipment (computers), we focus on the case in which there are two types of equipment: $\kappa_1 = \textit{computers}$ and $\kappa_2 = \textit{other equipment}$. In our baseline model in which a worker uses exactly one type of equipment, we construct $\pi_t(\lambda, \kappa, \sigma)$ for $\kappa = \textit{computers}$ as hours worked in occupation σ by those in labor type λ who respond that they use a computer at work relative to total hours worked by labor type λ in period t. Similarly, we construct $\pi_t(\lambda, \kappa_2, \sigma)$ as the hours worked in occupation σ by type λ workers who respond that they do not use a computer at work relative to total hours worked by labor type λ.

Constructing $\pi_t(\lambda, \kappa, \sigma)$ without using any information on non-computer equipment allocation introduces two related problems. In practice, workers who do not use computers may also not use other non-computer equipment, and workers who do use computers may also be more likely to use other non-computer equipment. Hence, we may incorrectly infer that computers have a comparative advantage in some occupations and that some worker types have a comparative advantage with computers. The German *Qualification and Working Conditions* survey circumvents some of these problems by asking whether or not a worker uses many different types of equipment and/or asks about the share of time that a worker uses computers. Specifically, one can use the 2006 survey question "How much of your total work time do you spend on computers?" to construct $\pi_t(\lambda, \kappa_1, \sigma)$ as the hours worked in occupation σ using computers by those in labor type λ relative to total hours worked by labor type λ and $\pi_t(\lambda, \kappa_2, \sigma)$ as the hours worked in occupation σ not using computers by those in labor type λ relative to total hours worked by labor type λ.[17] Constructing $\pi_{2006}(\lambda, \kappa, \sigma)$ in this way using the German data, we find similar patterns of comparative advantage as in the U.S. data.[18]

[17]DiNardo and Pischke (1997) is perhaps the best known paper using this survey. We discuss their critique of Krueger (1993) below.

[18]In the version of our model in which tasks are mapped to sectors in the data (see Appendix B) we can further compare our model's implications to U.S. data we do not use in the estimation. Specifically, the BEA reports data on the allocation of capital (aggregated across all workers) to sectors. We show in Appendix B that the model's implied allocation of computer and non-computer equipment across sectors matches the allocation that is observed in the data quite well.

5.2 Parameterization

Proposition 2 lists the parameters that we require to conduct our decomposition. As described above, we take $L_t(\lambda)/L_t(\lambda_1)$ for $t = t_0, t_1$ directly from the data. In this section we discuss how we assign values to the parameters ρ, θ, and α and how we estimate the values of the remaining parameters. Finally, combining the parameterization and theory, we show which components of our decomposition depend on observed changes in wages and which do not.

5.2.1 Assigned parameters

We assign the values of ρ, θ, and α as follows. In our baseline decomposition we set the elasticity of substitution between tasks, ρ, to 1. In our robustness section we show that whereas lowering ρ to 0.5 or raising it to 2 only modestly affects the importance of the combination of changes in labor supply, task shifters, and capital productivity relative to the importance of labor productivity, it does affect the importance of changes in task shifters relative to capital productivity. The parameter α determines the share of payments to equipment. When $\rho = 1$, one can show analytically that the value of $\alpha \in (0,1)$ does not impact any of our decomposition. In our robustness section, where we consider alternative values of ρ, we set $\alpha = 0.24$, consistent with the estimates in Burstein et al. (2013). The parameter θ determines the dispersion of idiosyncratic productivity draws. As discussed in Lagakos and Waugh (2013) and Hsieh et al. (2013) in models without capital, the dispersion of wages across workers within a labor group λ using capital κ in task σ obeys a Frechet distribution with shape parameter θ. In our baseline we set $\theta = 3.1$, which is in the mid-range of the estimates in Hsieh et al. (2013). In the robustness section, we consider a range of alternative values for θ.

5.2.2 Estimated parameters

In what follows we provide an overview of how we identify $\psi\tau(\lambda,\kappa,\sigma)$ for an arbitrary constant ψ, $\tau_{\lambda t}(\lambda)/\tau_{\lambda t}(\lambda_1)$, $\tau_{\sigma t}(\sigma)/\tau_{\sigma t}(\sigma_1)$, and $\tau_{\kappa t}(\kappa)/\tau_{\kappa t}(\kappa_1)$ for all λ, κ, σ and t (where t =1984, 1989, 1993, 1999, and 2003) . We provide details in Appendix A.3.

Our estimation procedure accounts for possible error in our observed measurement of $\{w_t(\lambda)\}$ and $\{\pi_t(\lambda,\kappa,\sigma)\}$: $\{w_t^*(\lambda)\}$ and $\{\pi_t^*(\lambda,\kappa,\sigma)\}$. The error terms $\{\iota_{1t}(\lambda,\kappa,\sigma)\}$ and $\{\iota_{2t}(\lambda)\}$ are due to sampling error:

$$\begin{aligned} \pi_t^*(\lambda,\kappa,\sigma) &= \pi_t(\lambda,\kappa,\sigma)\,\iota_{1t}(\lambda,\kappa,\sigma) \text{ for all } (\lambda,\kappa,\sigma) \text{ and } t \\ w_t^*(\lambda) &= w_t(\lambda)\,\iota_{2t}(\lambda) \text{ for all } \lambda \text{ and } t. \end{aligned} \tag{15}$$

This sampling error may be due to the individual-specific structural error ε—which induces workers to choose different tasks and equipment and to earn different wages—and to possible misreporting of equipment type, task, and wages by each worker. Because the error terms $\iota_{1t}(\lambda, \kappa, \sigma)$ and $\iota_{2t}(\lambda)$ are averages of errors affecting individual observations, they become arbitrarily close to one as the number of individuals sampled within each λ goes to infinity. This implies that our estimates described below are consistent for the true parameter values as the sample size per-λ goes to infinity.[19]

Our estimation involves three steps. In the first step we estimate the parameters that determine comparative advantage, $\tau(\lambda, \kappa, \sigma)$. In the second step we estimate equipment productivity and task shifters for each year, $\tau_{\kappa t}(\kappa)$ and $\tau_{\sigma t}(\sigma)$.[20] In the final step we estimate labor productivity for each year, $\tau_{\lambda t}(\lambda)$, using the estimates from steps one and two.

Step 1: Comparative advantage. Equation (14) gives us

$$\log \tau(\lambda, \kappa, \sigma) = \frac{1}{T}\sum_{t=1}^{T} \log \tau_t(\lambda, \kappa, \sigma) \tag{16}$$

for all λ, κ, σ. That is, given the definition of $\tau(\lambda, \kappa, \sigma)$ and $\tau_{tx}(x)$ for each $x = \lambda, \kappa, \sigma$, the log of $\tau(\lambda, \kappa, \sigma)$ is simply equal to the average across time of the log of $\tau_t(\lambda, \kappa, \sigma)$; see Appendix A.3. Equations (4) and (5) give the following relationship between $\tau_t(\lambda, \kappa, \sigma)$, wages, and allocations,

$$\tau_t(\lambda, \kappa, \sigma) = \gamma^{-\theta} w_t(\lambda)^{\theta} \pi_t(\lambda, \kappa, \sigma). \tag{17}$$

In Appendix A.3 we show that combining equations (15), (16), and (17) we obtain a consistent estimator of $\psi\tau(\lambda, \kappa, \sigma)$, where $\psi = \gamma^{\theta}$.

Step 2: Equipment productivity and task shifters. All else equal, a high value of $\tau_{\kappa t}(\kappa)/\tau_{\kappa t}(\kappa_1)$—which corresponds either to a low relative rental rate for κ or a high relative productivity of using κ—induces a large share of workers to use κ. Hence, we might expect to identify a high value of $\tau_{\kappa t}(\kappa)/\tau_{\kappa t}(\kappa_1)$ if the share of workers using κ is large. However, this intuition is incomplete. There are two other reasons this share may be large. First, there might be a large share of employment in tasks in which κ has a comparative advantage. Second, there might be a large supply of workers who have a comparative advantage using κ.

Our theory implies a clear strategy to identify $\tau_{\kappa t}(\kappa)/\tau_{\kappa t}(\kappa_1)$ that overcomes both of

[19] After dropping observations, in each year—1984, 1989, 1993, 1997, and 2003—we have between 52,000 and 62,000 workers in the October Supplement. With 30 labor types, this implies an average of about 1,900 observations per-λ in each year.

[20] Although step two does not require output from step one, it is pedagogically useful to separate them.

these concerns. Specifically, equation (14) gives us

$$\log \frac{\tau_{\kappa t}(\kappa)}{\tau_{\kappa t}(\kappa_1)} = \frac{1}{n_\Lambda n_\Sigma} \sum_{\lambda,\sigma} \log \frac{\tau_t(\lambda,\kappa,\sigma)}{\tau(\lambda,\kappa,\sigma)} \frac{\tau(\lambda,\kappa_1,\sigma)}{\tau_t(\lambda,\kappa_1,\sigma)}. \tag{18}$$

Hence, we identify a high value of $\tau_{\kappa t}(\kappa)/\tau_{\kappa t}(\kappa_1)$ in a given period if the share of workers using κ *within worker type and task pairs* (λ,σ) is relatively large. We estimate task fixed effects in a similar manner. Specifically, we use

$$\log \frac{\tau_{\sigma t}(\sigma)}{\tau_{\sigma t}(\sigma_1)} = \frac{1}{n_\Lambda n_K} \sum_{\lambda,\kappa} \log \frac{\tau_t(\lambda,\kappa,\sigma)}{\tau(\lambda,\kappa,\sigma)} \frac{\tau(\lambda,\kappa,\sigma_1)}{\tau_t(\lambda,\kappa,\sigma_1)}, \tag{19}$$

and identify a large value of $\tau_{\sigma t}(\sigma)/\tau_{\sigma t}(\sigma_1)$ in a given period if the share of workers employed in σ *within worker type and equipment pairs* (λ,κ) is relatively large. In Appendix A.3 we use the previous expressions, together with equations (15), (16), and (17) to obtain consistent estimators of $\tau_{\kappa t}(\kappa)/\tau_{\kappa t}(\kappa_1)$ and $\tau_{\sigma t}(\sigma)/\tau_{\sigma t}(\sigma_1)$.

Step 3: Labor productivity. All else equal, a higher value of $\tau_{\lambda t}(\lambda)/\tau_{\lambda t}(\lambda_1)$ raises the relative wage of λ. However, as in step 2, observing a higher relative wage for λ does not necessarily imply a higher relative value of $\tau_{\lambda t}(\lambda)$ for two reasons. First, λ workers would earn relatively more if the tasks in which they have a comparative advantage had a larger task shifter, $\tau_{\sigma t}(\sigma)$. Second, they would earn relatively more if the equipment with which they have a comparative advantage were relatively more productive, $\tau_{\kappa t}(\kappa)$. Again, our theory implies a clear strategy to identify $\tau_{\lambda t}(\lambda)/\tau_{\lambda t}(\lambda_1)$. Equations (5) and (14) give us

$$\log \frac{\tau_{\lambda t}(\lambda)}{\tau_{\lambda t}(\lambda_1)} = \theta \log \frac{w_t(\lambda)}{w_t(\lambda_1)} - \log \frac{\sum_{\kappa,\sigma} \tau_{\kappa t}(\kappa)\,\tau_{\sigma t}(\sigma)\,\tau(\lambda,\kappa,\sigma)}{\sum_{\kappa',\sigma'} \tau_{\kappa t}(\kappa')\,\tau_{\sigma t}(\sigma')\,\tau(\lambda_1,\kappa',\sigma')}. \tag{20}$$

Equation (20) identifies relative worker productivities to exactly match relative wages, controling for worker comparative advantage, equipment productivities, and task shifters. In Appendix A.3 we show that the previous expression, together with equations (15), (16), and (17) and the consistent estimates of $\tau_{\sigma t}(\sigma)/\tau_{\sigma t}(\sigma_1)$ and $\tau_{\kappa t}(\kappa)/\tau_{\kappa t}(\kappa_1)$ yield a consistent estimator of $\tau_{\lambda t}(\lambda)/\tau_{\lambda t}(\lambda_1)$.[21]

Equation (16), ((18)), and (19) assume that $\pi_t(\lambda,\kappa,\sigma)$ is larger than 0 for every time period t and every triplet (λ,κ,σ). In the data, some of the observed values of $\pi_t(\lambda,\kappa,\sigma)$ are actually 0.[22] For the main results presented in the paper, the sample averages in

[21] Appendix A.3 also describes an alternative strategy that yields similar conclusions.

[22] In particular, the fraction of observations that are equal to 0 in the data are: 17%, for 1984, 13% in 1989,

equations (16), (18), and (19) are computed using data on only the positive $\pi_t(\lambda, \kappa, \sigma)$.[23]

5.2.3 The role of data on wage changes in the decomposition

How do observed relative wage changes between t_0 and t_1 shape the results of our decomposition? Here we show that, given the estimation strategy introduced in the previous section, observed wages in period t_1 do not affect the labor supply or capital productivity components; moreover, conditional on total labor income earned in each task, they do not affect our task shifters component either. This is in contrast to the labor productivity component, which as can be seen in equation (20) depends directly on observed changes in wages.

The labor supply component of our decomposition is clearly independent of wages in period t_1, since the only data from period t_1 that enters the algorithm in Lemma 1 is $L_{t_1}(\lambda)$, which we assume is exogenous. Our capital productivity component is also independent of wages in period t_1, since the only data from period t_1 that enters the algorithm in Lemma 3 is $\hat{\tau}_\kappa(\kappa) / \hat{\tau}_\kappa(\kappa_1)$, and our estimate of $\tau_{\kappa t}(\kappa) / \tau_{\kappa t}(\kappa')$ is independent of wages, as shown in Appendix A.3. Finally, according to the algorithm presented in Lemma 4, our task shifters component requires $\tau_{\sigma t_1}(\sigma) / \tau_{\sigma t_1}(\sigma_1)$, which is independent of wages in period t_1 as shown in Appendix A.3, as well as the share of total labor income in each task in period t_1, which does depend on wages in period t_1 if $\rho \neq 1$. Hence, in general when $\rho \neq 1$, the task shifters component is independent of wages in period t_1 only conditional on total labor income in each task in period t_1. In the special case in which $\rho = 1$, the task shifters component is independent of wages in period t_1.

Since changes in wages are not an input in the algorithm to determine the direct contribution of these components, there is no *a priori* reason to believe that the labor supply, task shifter, or capital productivity exercises should play an important role in explaining either the qualitative or quantitative pattern of relative wages over time. The opposite, however, is true of the labor productivity component, since relative labor productivities are estimated to match relative wages.

1993, and 1997, and 8% in 2003.

[23]For robustness, we have redone our estimation and decomposition using a higher degree of worker-level aggregation. For instance, with only five worker types—the five education levels—there are 1.5% missing values for 1984, 1989, and 1993 and 0.5% missing values for 1997 and 2003. The results we obtain under this aggregated definition of worker types are similar to those obtained in our baseline, suggesting that missing values are not important for generating our results.

5.3 Estimation results

In this section we present summary statistics describing our estimated parameters.

5.3.1 Patterns of comparative advantage

Our estimation procedure recovers estimates of the comparative advantage parameters, $\tau(\lambda,\kappa,\sigma)$, without imposing any functional form restriction on how these vary across (λ,κ,σ). Solely for the purpose of summaring the data, we project our estimated values of $\tau(\lambda,\kappa,\sigma)$ onto some observable characteristics of workers and tasks as well as their corresponding interaction terms. We restrict $\tau(\lambda,\kappa,\sigma) = \tau(\lambda,\sigma)\,\tau(\lambda,\kappa)\,\tau(\kappa,\sigma)$—in which case comparative advantage between (i) workers and equipment is common across tasks, (ii) workers and tasks is common across equipment, and (iii) equipment and tasks is common across workers. Specifically, we impose

$$\tau(\lambda,\kappa,\sigma) = \exp\left(\sum_{i=1}^{n_\lambda}\sum_{j=1}^{n_\sigma}\beta_{ij}X_i(\lambda)\,X_j(\sigma)\right)\exp\left(\sum_{i=1}^{n_\lambda}\beta_{\lambda i}(\kappa)\,X_i(\lambda)\right)\exp\left(\sum_{j=1}^{n_\sigma}\beta_{\sigma j}(\kappa)\,X_j(\sigma)\right) \quad (21)$$

where $X(\lambda) \geq 0$ and $X(\sigma) \geq 0$ are vectors of n_λ and n_σ worker and task characteristics described below (which are distinct from the number of worker types, n_Λ, and tasks, n_Σ); β is a vector with $n_\lambda n_\sigma$ elements; and $\beta_\lambda(\kappa)$ and $\beta_\sigma(\kappa)$ are vectors with n_λ and n_σ elements, respectively, where there is one $\beta_\lambda(\kappa)$ and one $\beta_\sigma(\kappa)$ for each type of equipment κ.

The vectors β, $\beta_\lambda(\kappa)$, and $\beta_\sigma(\kappa)$ summarize comparative advantage. According to equation (21), $\beta_{ij} > 0$ implies that a high value of worker characteristic i (e.g., education) is relatively more productive when employed in a task characterized by a high value of characteristic j (e.g., the importance of analyzing data and information). Relatedly, $\beta_{\lambda i}(\kappa) - \beta_{\lambda i}(\kappa') > 0$ implies that a high value of worker characteristic i (e.g., education) is relatively more productive when using equipment κ (e.g. computers) than κ' (e.g. non-computers). Finally, $\beta_{\sigma j}(\kappa) - \beta_{\sigma j}(\kappa') > 0$ implies that equipment κ (e.g. computers) is relatively more productive in tasks characterized by a high value of characteristic j (e.g., the importance of repeating the same task) relative to κ' (e.g. non-computers).

We include three worker characteristics, constructed using the March CPS: age, gender, and education. We measure age and education in years, as the average within λ across all t. Gender is an indicator function that equals one if λ corresponds to a female labor type. We use seven task characteristics, which we measure by merging job task requirements from O*NET to their corresponding Census occupation classifications, following Acemoglu and Autor (2011). We provide details in Appendix A.2. We use

$X(\lambda)$	$\beta_\lambda(\kappa_1) - \beta_\lambda(\kappa_2)$	$X(\sigma)$	$\beta_\sigma(\kappa_1) - \beta_\sigma(\kappa_2)$	$\beta_{education \times j}$
Age	-0.003	Analyzing data/information	0.180	0.269^a
Female	0.237^b	Thinking creatively	0.499^a	0.033
Education	0.330^a	Guiding, directing, motivating	0.001	-0.138^a
		Importance of repetition	0.570^a	-0.128^a
		Pace determined equipment	-0.474^a	-0.028
		Manual dexterity	-0.628^a	-0.086^a
		Social Perceptiveness	-0.728^a	0.088^c

Table 2: Comparative advantage

a, b, and c denote significance at the 99%, 95%, and 90% levels, where standard errors are robust to heteroskedasticity

the following 7 O*NET scales, each of which is between zero and ten: (i) Analyzing data/information; (ii) Thinking creatively; (iii) Guiding, directing, and motivating subordinates; (iv) Importance of repeating the same tasks; (v) Pace determined by speed of equipment; (vi) Manual dexterity; and (vii) Social Perceptiveness.[24] Table 10 in Appendix A.2 lists these scales for each of the twenty occupations.

In Appendix A.3 we show how to estimate β, $\beta_\lambda(\kappa)$, and $\beta_\sigma(\kappa)$. Here we use these estimates to summarize patterns of comparative advantage. Table 2 lists the parameter vectors $\beta_\lambda(\kappa)$ and $\beta_\sigma(\kappa)$ as well as the components of β_{ij} that refer to $i =$ education. This table highlights three important results. First, each year of additional education raises productivity in computer relative to non-computer equipment. Given two workers of the same age and gender employed in the same occupation and with the same idiosyncratic component of productivity, the one with a college degree (16 years of education) is about $\exp(0.33 \times 4/\theta) = 1.53$ times more productive with computers (relative to non-computer equipment) than the one with a high school degree (12 years of education). Second, computers are relatively productive in tasks in which thinking creatively and repetition are relatively important and relatively unproductive in tasks in which the pace is determined by equipment and in which manual dexterity and social perceptiveness are relatively important. Finally, more educated workers have a comparative advantage in tasks in which analyzing data/information is relatively important and have a comparative disadvantage in tasks in which guiding, directing, and motivating subordinates, repetition, and manual dexterity are relatively important.

[24] To provide some context for these scales, Acemoglu and Autor (2011) incorporate (i) and (ii) into their measure of "Non-routine cognitive: Analytical," (iii) into "Non-routine cognitive: Interpersonal," (iv) into "Routine cognitive," (v) into "Routine manual," and (vi) into "Non-routine manual physical."

	84-89	89-93	93-97	97-03	**84-03**
Capital					
κ_1/κ_2	0.52	0.50	0.11	0.38	**1.51**
Education					
HSG/HSD	0.00	-0.06	0.02	-0.03	**-0.05**
SMC/HSD	0.03	-0.12	0.03	-0.09	**-0.11**
CLG/HSD	0.07	-0.06	0.06	0.03	**0.12**
GTC/HSD	0.09	0.15	0.14	0.00	**0.40**

Table 3: Changes over time in log relative capital and labor productivities.
HSD: high school dropout; HSG: high school graduate; SMC: some college; CLG: college; GTC: graduate training

5.3.2 Changes in task shifters, labor productivities, equipment productivities

Table 3 summarizes changes in capital equipment productivity. The relative productivity of computer capital rises between each pair of years in our sample. This rise in the model corresponds with the extraordinarily large increase in the quantity and decrease in the price of computer equipment relative to all other capital equipment and relative to structures capital measured by the BEA (which we do not use in the estimation), as reviewed in the introduction.

Table 3 also summarizes changes in worker productivity. We aggregate up from thirty labor groups to five education groups and display changes in productivity of each group relative to the lowest education group: high school dropouts. Note that over the full sample, changes in worker productivity are non-monotonic—intermediate education levels become relatively less productive than both low and high education levels—and this non-monotonicity is driven by changes ocurring after 1989.

Finally, we find that task shifters rose in tasks in which educated workers have a comparative advantage and shrank in tasks in which they have a comparative disadvantage. Table 10 in Appendix A.4 reports the change in task shifters for each occupation between 1984 and 2003; for example, the largest and smallest task shifters are in the "health services" and "machine operators, assemblers, inspectors" occupations, respectively. As discussed above, estimated task shifters are shaped in general both by estimated $\hat{\tau}_\sigma(\sigma)$ and estimated changes in incomes across tasks, $\hat{\zeta}(\sigma)$. As shown in Lemma 4 in Appendix A.1, task shifters are given by $\frac{\hat{\zeta}(\sigma)}{\hat{\zeta}(\sigma_1)}\left(\frac{\hat{\tau}_\sigma(\sigma)}{\hat{\tau}_\sigma(\sigma_1)}\right)^{(\rho-1)(1-\alpha)/\theta}$. Note that when the aggregate production function is Cobb Douglas, $\rho = 1$, only changes in preference parameters, $\hat{\mu}(\sigma)$, matter for the task shifter counterfactuals on wages; and these changes are fully summarized by estimated changes in incomes across tasks $\hat{\zeta}(\sigma)/\hat{\zeta}(\sigma_1)$.[25] Regressing $\hat{\zeta}(\sigma)/\hat{\zeta}(\sigma_1)$ (be-

[25]When $\rho = 1$, changes in task-level productivities, $\hat{T}_\sigma(\sigma)$, are irrelevant for relative wages because task prices and productivities adjust proportionately.

	Data	Labor supply	Task shifters	Labor prod.	Capital prod.
84-89	0.054	-0.020	0.025	0.017	0.030
89-93	0.062	-0.009	0.022	0.020	0.029
93-97	0.025	-0.014	0.015	0.017	0.006
97-03	0.046	-0.025	0.032	0.020	0.020
84-03	**0.182**	**-0.070**	**0.096**	**0.070**	**0.079**

Table 4: Decomposing changes in the log skill premium

tween 1984 and 2003) separately on each of the seven task characteristics derived from O*NET and discussed above yields three significant coefficients (each is significant at the 1% level). Occupations in which the pace is particularly determined by equipment and in which manual dexterity is particularly important (occupations in which educated workers have a comparative disadvantage according to Table 2) shrank whereas occupations in which social perceptiveness is particularly important (occupations in which educated workers have a comparative advantage according to Table 2) grew.

6 Results

Combining our parameterization and theory, we now turn to the results of our decomposition.

Skill premium. We begin by decomposing changes in the composition-adjusted skill premium between each pair of consecutive years and over the full sample, displayed in Table 4. The first column reports the change in the data, which is also the change predicted by the model when all changes (in labor supply, task shifters, labor productivity, and capital productivity) are simultaneously considered. The subsequent four columns summarize the change in the skill premium predicted by the model for each component of the decomposition separately. While the sum of changes in the skill premium predicted by the four components need not sum to the total predicted change in the skill premium due to interactions, in practice the difference is very small.

Over the full sample, between 1984 and 2003, changes in capital productivity and task shifters are the most important forces driving changes in the skill premium; see the final row of Table 4. The capital productivity component alone accounts for roughly 43% of the rise in the skill premium ($0.43 \simeq 0.079/0.182$) and roughly 32% of the sum of the forces pushing the skill premium upwards ($0.32 \simeq 0.079/(0.096 + 0.070 + 0.079)$). Over subperiods, changes in capital productivity are particularly important in generating changes in the skill premium over 1984-1989, 1989-1993, and 1997-2003. These are precisely the years in which the overall share of workers using computers rose most rapidly; see Table 1. We obtain the result that computerization has substantially increased the U.S. skill premium because we find: (*i*) strong education-computer comparative advantage (see Table 2), (*ii*) a substantial share of workers using computers (see Table 1), and (*iii*) large growth in computer usage within worker-task pairs (see Table 3).[26]

The task shifter component is even more important than the capital productivity component over the full sample. It accounts for rougly 53% of the rise in the skill premium and 39% of the sum of the forces pushing the skill premium upwards. We obtain the result that task shifters have substantially increased the U.S. skill premium (given our choice of parameters) because we find: (*i*) strong education-occupation comparative advantage, (*ii*) a substantial share of workers in the expanding or contracting occupations, and (*iii*) large changes in task shifters.

Perhaps surprisingly, of the mechanisms pushing the skill premium upwards over the full sample, the weakest is the one mechanism that was estimated to match observed relative wages (and, therefore, changes in relative wages): labor productivity (recall our discussion in Section **5.2.3**). Over the full sample this component accounts for roughly 38% of the rise in the skill premium and roughly 29% of the sum of the forces pushing the skill premium upwards.

Disaggregated groups. Table 5 decomposes changes in between-education-group wage inequality at a higher level of disaggregation, comparing changes in composition-adjusted average wages across the five education groups over the full sample, 1984-2003. The results reported in Table 4 are robust: the labor productivity component is not particularly important for explaining *the rise* in between education-group inequality even at this more disaggregated level. It either pushes the relative wage of education groups in the wrong direction or accounts for a relatively small share of the rise in the relative wage of more educated workers.

However, Table 5 demonstrates that, over the full sample, the impact of changes in labor productivity on relative wages across education groups are *U-shaped*: they decrease

[26] DiNardo and Pischke (1997) critique Krueger (1993) by showing that pencils can explain wage premia as well as computers. Their critique does not apply here for two reasons. First, our approach is fundamentally different from Krueger (1993). Instead of using the October Supplement to regress wages on computer usage, we use it to identify comparative advantage. Second, in order for pencils to drive changes in wages (as we find computers do), we would have to find (*i*) strong worker-pencil comparative advantage (identified within occupations), (*ii*) a large share of workers using pencils, and (*iii*) extremely large and systematic changes in pencil usage within worker-task pairs over time. Given the extraordinary decline (rise) over time in the relative price (quantity) of computer equipment compared to all other equipment and structures, this is not a reasonable concern.

	Data	Labor supply	Task shifters	Labor prod.	Capital prod.
HSG/HSD	0.035	-0.024	0.040	-0.019	0.043
SMC/HSD	0.068	-0.052	0.083	-0.045	0.085
CLG/HSD	0.182	-0.091	0.128	0.015	0.128
GTC/HSD	0.286	-0.113	0.169	0.096	0.129

Table 5: Decomposing changes in log relative wages between education groups: 1984-2003
HSD: high school dropout; HSG: high school graduate; SMC: some college; CLG: college; GTC: graduate training

wages of intermediate education groups relative to the least educated group and relative to the most educated groups. Table 3 provides the intuition for this result: labor productivity was estimated to rise in the extreme education groups relative to the intermediate ones. Hence, whereas changes in labor productivity are not the most important force driving the rise in between-education group inequality, at a disaggregated level they do play an important role: they generate "wage polarization," a feature of changes in wage distributions in a number of countries over the last few decades; see e.g. Autor et al. (2008) and Goos et al. (2009).

To further document this result, Figure 1 plots a cubic fit of the log change in average hourly wages between 1989-2003 for the 15 male labor types against the log of the average hourly wage in 1989. Even with only 15 labor types, we observe the wage polarization that others have documented in the full (and especially male) income distribution following 1988 in the U.S.; see e.g. Autor et al. (2008) and Acemoglu and Autor (2011). This figure also plots the log change in average hourly wages between 1989-2003 predicted by the model from the combination of the labor supply, task shifter, and capital productivity components. These changes do not generate wage polarization, either individually or when combined. Instead, wage polarization is accounted for by changes in labor productivity, as shown in Figure 1.[27]

Gender. Between 1984-2003 the log change in the composition adjusted gender wage gap (the average wage of males relative to females) in the data was -0.125. According to our decomposition, the rise in female labor supply increased the gender wage gap by 0.024 log points, task shifters decreased it by 0.067 log points, labor productivity decreased it by 0.071 log points, and capital productivity decreased it by 0.005 log points. These numbers highlight three important results. First, changes in task shifters and labor productivity are the most imortant forces driving changes in the gender wage gap between 1984-2003;

[27] A similar conclusion emerges from the other periods in which we observe wage polarization in the data, when we use a quadratic rather than cubic fit, and if we match tasks in the model to sectors in the data.

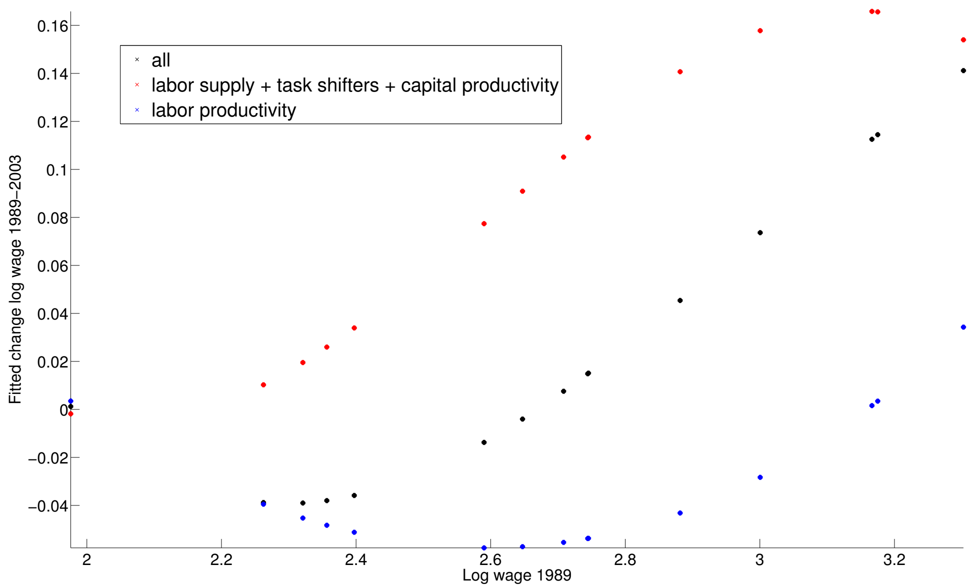

Figure 1: Cubic fit of the log change in average hourly wages between 1989-2003 (relative to the lowest wage group in 1989) plotted against log average hourly wages in 1989.

together they account for roughly 97% of the forces decreasing the gender wage gap. Second, they are almost equally important over this time period. Finally, capital productivity has almost no effect on the gender wage gap. It is crucially important for this final result that we estimate worker-computer comparative advantage using allocations to computers within occupations rather than at the aggregate level. In spite of the fact that women are substantially more likely to use a computer at work than men—see Table 1—much of the difference in computer usage across genders is accounted for by differences in the occupations to which men and women are allocated, rather than by differences in computer usage within occupations.

7 Robustness and sensitivity analysis

In this section we consider two types of sensitivity exercises. First, we illustrate the importance of all three types of comparative advantage by turning some of them off. Second, we perform sensitivity to different values of ρ and θ.

	Labor supply	Task shifters	Labor prod.	Capital prod.
Baseline	-0.070	0.096	0.070	0.079
Only labor-equipment CA	0	0	0.013	0.169
Only labor-task CA	-0.075	0.108	0.142	0

Table 6: Decomposing changes in the log skill premium under different assumptions on comparative advantage: 1984-2003

7.1 Sources of comparative advantage

Our model features three types of comparative advantage: *(i)* between labor and equipment, *(ii)* between equipment and tasks, and *(iii)* between labor and tasks. To demonstrate the importance of including each of these various sources of comparative advantage, we perform two exercises. First, we abstract from comparative advantage related to equipment, i.e. *(i)* and *(ii)*. To do so, we impose in our estimation that $\tau(\lambda, \kappa_1, \sigma) = \tau(\lambda, \kappa_2, \sigma)$. This is equivalent, in terms of the model's implications for changes in relative wages, to assuming that there a single equipment good. Second, we abstract from comparative advantage related to tasks, i.e. *(ii)* and *(iii)*, imposing in our estimation that $\tau(\lambda, \kappa, \sigma_i) = \tau(\lambda, \kappa, \sigma_1)$ for $i = 2, ..., n_\Sigma$. This is equivalent—again in terms of changes in relative wages—to assuming that there is a single task.

Abstracting from any comparative advantage at the level of tasks (i.e. assuming away worker-task and equipment-task comparative advantage) has two effects. First, it implies that the labor supply and task shifters components of our decomposition go to zero. Since the sum of these components pushes the skill premium up, this implies that holding fixed the importance of the capital productivity component, the labor productivity component must increase. Second, it implies that the only force generating the allocation of worker types to equipment types is direct comparative advantage. Since we found in our baseline exercise that educated workers use computers relatively more than non-computer equipment both because of direct and indirect comparative advantage, abstracting from any comparative advantage at the level of tasks magnifies the stength of worker-equipment comparative advantage and increases the impact of the capital productivity component, thereby reducing the impact of the labor productivity component. Table 6 confirms this intuition: the strength of the capital productivity component becomes much stronger in the absence of any comparative advantage at the level of tasks, so much so that the labor productivity component becomes weaker. Abstracting from any comparative advantage at the level of tasks, we would incorrectly conclude that almost all of the rise in the skill premium (93%) has been driven by changes in relative equipment productivities.

Abstracting from any comparative advantage at the level of equipment has similar effects. First, it implies that capital productivity component of our decomposition goes to zero so that—holding fixed the importance of the labor supply and task shifters components—the absolute value of the labor productivity component must become larger. Second, it implies that the only force generating the allocation of worker types to tasks is the direct comparative advantage. Since we found in our baseline exercise that educated workers are employed in expanding tasks both because of direct and indirect comparative advantage, abstracting from any comparative advantage at the level of equipment magnifies the stength of this direct comparative advantage. Table 6 shows that abstracting from any comparative advantage at the level of equipment magnifies the importance of labor productivity in explaining the rise of the skill premium.

In summary, abstracting from any comparative advantage at the level of either tasks or equipment has a large impact on the composition of changes in between-group inequality. It does so first by forcing changes in labor productivity to absorb the impacts of the missing component(s) and second by changing the inferred strength of the remaining source of comparative advantage.

7.2 Alternative parameter values

In this section we vary θ and ρ—recall that our decomposition results are indepenent of the value of $\alpha \in (0,1)$ given our baseline value $\rho = 1$—and report the implications of these alternative values for our decomposition. We focus on changes only in the skill premium and only over the full sample.

Alternative values for θ. A higher value of θ corresponds to less dispersion in idiosyncratic productivities, ε, and—as shown in equation (12)—increases the elasticity of worker allocation, π, with respect to changes in prices, rental rates, and productivities. Hence, the same change in underlying primitives yields smaller changes in average wages. Accordingly, a higher values of θ will reduce the impact of changes in labor supply, task shifters, and capital productivity on relative wages. In response, changes in labor productivity must contribute more to the rise in between-education group inequality, since $\tau_\lambda(\lambda) / \tau_\lambda(\lambda_1)$ is estimated to match relative wages.

Table 7 confirms this intuition. The middle row of the table replicates our baseline results. We consider two extreme values of θ, $\theta = 2$ and $\theta = 4$, in addition to two alternative values that are consistent with estimates from Hsieh et al. (2013), $\theta = 2.9$ and $\theta = 3.3$. Table 7 demonstrates that our baseline result—that the combination of changes in task shifters and capital productivity explains the majority of the rise in between-education

value of θ	Labor supply	Task shifters	Labor prod.	Capital prod.
$\theta = 2$	-0.099	0.130	0.032	0.110
$\theta = 2.9$	-0.074	0.101	0.065	0.083
$\theta = 3.1$	-0.070	0.096	0.070	0.079
$\theta = 3.3$	-0.066	0.092	0.075	0.075
$\theta = 4$	-0.056	0.080	0.089	0.064

Table 7: Decomposing changes in the skill premium for alternative values of θ: 1984-2003

group inequality—is largely robust to alterative values of θ within the range of $2 \leq \theta \leq 4$. At $\theta = 2$ labor productivity plays almost no role in increasing the skill premium between 1984 and 2003 whereas at $\theta = 4$ it becomes relatively more important, but continues to explain substantially less of the rise in the skill premium than the combination of task shifters and capital productivity.

Alternative values for ρ. ρ is the elasticity of substitution across tasks in the aggregate production function. The impact of ρ on the effect of changes in labor supply is straightforward. An increase in the relative supply of a given worker type tends to depress the prices of the tasks in which that worker type has a comparative advantage, thus decreasing that worker type's relative wage and the relative wages of workers who have similar patterns of comparative advantage across tasks. The larger is ρ, the less responsive are relative task prices and the weaker is this effect. The impact of ρ on the direct effect of changes in capital productivity is related, and was discussed in Section 3.4: a larger value of ρ raises the impact, in response to changes in capital productivity, of changes in rental rates relative to task prices.

The intuition for the impact of ρ on task shifters is more complicated. Between 1984 and 2003, we observe (i) an increase in income in skill-intensive occupations (those occupations in which educated workers are disproportionately allocated) and (ii) changes in labor supply and capital productivity that tend to decrease the relative prices of these occupations. If ρ is low, (ii) generates a large reduction in income in skill-intensive tasks. Hence, to match (i), a lower value of ρ requires larger task shifters in favor of skill-intensive tasks. Mechanically, our estimates of $\hat{\tau}_\sigma(\sigma)$, which are lower for skill-intensive occupations, are independent of ρ. Since what matters for the impact of task shifters is $\frac{\hat{\zeta}(\sigma)}{\hat{\zeta}(\sigma_1)}\left(\frac{\hat{\tau}_\sigma(\sigma)}{\hat{\tau}_\sigma(\sigma_1)}\right)^{(\rho-1)(1-\alpha)/\theta}$, a lower value of ρ is similar to increasing task shifters in skill-intensive occupations.

Table 8 confirms this intuition. The middle row of the table replicates our baseline results. Our baseline result—that the combination of changes in task shifters and capital

value of ρ	Labor supply	Task shifters	Labor prod.	Capital prod.
$\rho = 1/2$	-0.085	0.130	0.065	0.059
$\rho = 2/3$	-0.079	0.118	0.067	0.066
$\rho = 1$	-0.070	0.096	0.070	0.079
$\rho = 3/2$	-0.060	0.072	0.073	0.092
$\rho = 2$	-0.052	0.054	0.076	0.102

Table 8: Decomposing changes in the skill premium for alternative values of ρ: 1984-2003

productivity explains the majority of the rise in between-education group inequality—is robust to alternative values of ρ within the range of $1/2 \leq \rho \leq 2$: the importance of labor productivity remains stable. However, the relative importance of task shifters and capital productivity changes dramatically as we vary ρ from 1/2 to 2. At $\rho = 1/2$ task shifters are the dominant force explaining changes in the skill premium whereas at $\rho = 2$ capital productivity is.

8 International trade

Our theoretical and quantitative analyses have focused on a closed economy. In this section we extend our model to incorporate and quantify the impact of international trade on inequality. We assume that consumption and investment goods are traded whereas we abstract from trade in tasks (given the lack of data on trade in occupational output). Hence, in this section international trade only affects relative wages through relative prices of capital equipment goods and, therefore, equipment rental rates.

8.1 Environment and equilibrium

Environment. We denote countries by n. The final good is produced combining domestically performed tasks, as in equation (1). The output of this final good is used to produce country n's consumption good and country n's capital goods, satisfying the resource constraint given by

$$Y_{n,t} = q_{n,t}(C)\, Y_{n,t}(C) + \sum_{\kappa} q_{n,t}(\kappa)\, Y_{n,t}(\kappa).$$

Country n's consumption is a CES aggregator over consumption goods from all source countries,

$$C_{n,t} = \left(\sum_{i} C_{in,t}^{\frac{\eta(C)-1}{\eta(C)}} \right)^{\frac{\eta(C)}{\eta(C)-1}},$$

where $C_{in,t} \geq 0$ is consumption in country n of country i's good at time t. World market clearing in consumption goods requires

$$Y_{n,t}(C) = \sum_i d_{ni,t}(C) C_{ni,t}$$

where $d_{ni,t}(C) \geq 1$ is the iceberg trade cost for consumption goods from source country n to destination country i at time t. Similarly, country n's investment in κ is a CES aggregator over investment goods from all source countries,

$$I_{n,t}(\kappa) = \left(\sum_n I_{in,t}(\kappa)^{\frac{\eta(\kappa)-1}{\eta(\kappa)}} \right)^{\frac{\eta(\kappa)}{\eta(\kappa)-1}}$$

where $I_{in,t}(\kappa) \geq 0$ is country n's investment in country i's κ good at time t. World market clearing in investment κ goods requires

$$Y_{n,t}(\kappa) = \sum_i d_{ni,t}(\kappa) I_{ni,t}(\kappa)$$

where $d_{ni,t}(\kappa) \geq 1$ is the iceberg trade cost for investment good κ. Finally, the law of motion for capital κ is

$$K_{n,t+1}(\kappa) = (1 - dep_n(\kappa)) K_{n,t}(\kappa) + I_{n,t}(\kappa)$$

and utility of the representative household is given by $\sum_{t=0}^{\infty} u_{n,t}(C_{n,t})$. We assume that there are no intra-national trade costs: $d_{nn,t}(C) = d_{nn,t}(\kappa) = 1$ for all n, t, and κ. Note that this model reduces to our baseline model when countries are in autarky: $d_{ni,t}(C) = d_{ni,t}(\kappa) = \infty$ for all $n \neq i$, t, and κ.

Equilibrium in changes. Relative to the baseline closed-economy model summarized by equations (9)-(13), the only change is to equation (13). Along a balanced growth path, we now have

$$\hat{r}_n(\kappa) = \hat{P}_n(\kappa) = \hat{s}_{nn}(\kappa)^{1/(\eta(\kappa)-1)} \hat{P}_{nn}(\kappa)$$

and

$$\frac{\hat{P}_{nn}(\kappa)}{\hat{P}_{nn}(\kappa')} = \frac{\hat{q}_n(\kappa)}{\hat{q}_n(\kappa')}.$$

Here $s_{nn,t}(\kappa) = \frac{P_{nn,t}(\kappa) I_{nn,t}(\kappa)}{\sum_i P_{in,t}(\kappa) I_{in,t}(\kappa)}$ denotes expenditure on domestic investment good κ relative to total expenditure on investment good κ in country n (the "domestic aborption share"), $P_{in,t}(\kappa)$ denotes the price of country i's investment good in country n (inclusive of trade costs), and $P_n(\kappa)$ denotes the price of the aggregate investment good κ in country n (a CES aggregator of $P_{in,t}(\kappa)$ across i). The domestic absorption share is determined in the world general equilibrium. If country n's trade costs are set to infinity (i.e. $d_{in,t}(.) = \infty$ for $i \neq n$), then $s_{nn,t}(.) = 1$. In the counterfactual exercises described below, we consider either changing trade costs to infinity or matching observed changes over time in $s_{nn,t}(\kappa)$. Therefore, we do not need to specify conditions on trade balance or solve for the equilibrium determination of $s_{nn,t}(\kappa)$ in the world general equilibrium. Combining the two previous equations, we obtain

$$\frac{\hat{r}_n(\kappa)}{\hat{r}_n(\kappa')} = \frac{\hat{q}_n(\kappa)}{\hat{q}_n(\kappa')} \times \frac{\hat{s}_{nn}(\kappa)^{1/(\eta(\kappa)-1)}}{\hat{s}_{nn}(\kappa')^{1/(\eta(\kappa')-1)}} \tag{22}$$

Because task markets are autarkic, trade only affects relative wages through its impact on relative capital prices. Hence, given relative capital prices in country n, the equilibrium allocation of factors and relative wages in country n are determined exactly as in our baseline model.

The result that the effects of trade on allocations and prices can be summarized by changes in domestic absorption shares, $\hat{s}_{nn}(\kappa)$, and the gravity elasticity, $\eta(\kappa) - 1$, holds across a wide range of quantitative trade models; see Arkolakis et al. (2012). We assume an Armington trade model only for expositional simplicity.

With international trade in tasks, from which we have abstracted, trade would also potentially explain a portion of the task shifter component of our decomposition. Because the impact on relative wages of changes in task shifters is large in our decomposition, this suggests that the role of trade on wages through this channel could be substantial.

8.2 Counterfactual exercises

In this section we show how to connect our extended model to the data, provide two results that allow us to conduct counterfactuals, and quantify the impact of international trade on relative wages in the United States.

Connecting model to data. Because the equilibrium allocation of factors and relative wages are determined exactly as in our baseline model, for given rental rates and task prices, our estimating equations and procedure are unchanged relative to the baseline model. Whereas the definitions of all estimated parameters is the same as in the baseline model, changes in estimated capital productivities $\hat{\tau}_{n,\kappa}(\kappa)$ now capture changes in domestic technologies as in our baseline model as well as changes in all international technologies, factor supplies, and trade costs, as summarized by changes in domestic absorption shares $\hat{s}_{nn}(\kappa)$,

$$\frac{\hat{\tau}_{n,\kappa}(\kappa_1)}{\hat{\tau}_{n,\kappa}(\kappa_2)} = \left(\frac{\hat{T}_{n,\kappa}(\kappa_1)}{\hat{T}_{n,\kappa}(\kappa_2)}\left(\frac{\hat{q}_n(\kappa_1)}{\hat{q}_n(\kappa_2)}\right)^{\frac{-\alpha}{1-\alpha}}\right)^{\theta}\left(\frac{\hat{s}_{nn}(\kappa_2)^{1/(\eta(\kappa_2)-1)}}{\hat{s}_{nn}(\kappa_1)^{1/(\eta(\kappa_1)-1)}}\right)^{\frac{\theta\alpha}{1-\alpha}} \tag{23}$$

Counterfactuals. We use our framework to conduct two counterfactual exercises quantifying the impact of international trade on relative wages through its impact on the relative price of capital equipment. In the first counterfactual we hold all parameters in country n fixed and increase trade costs between country n and its trade partners such that country n moves to autarky. This counterfactual quantifies the impact on wages in country n if it were to move to autarky at time t, holding all of country n's paramaters fixed, which we denote by $\hat{w}^A_{n,t}(\lambda)$. The counterfactual change in the wage of λ workers relative to λ' workers is $\hat{w}^A_{n,t}(\lambda)/\hat{w}^A_{n,t}(\lambda')$. Conducting this counterfactual is straightforward given equation (22) and is summarized by the following proposition.

Proposition 3. *$\hat{w}^A_{n,t}(\lambda)/\hat{w}^A_{n,t}(\lambda_1)$ for all λ is the solution to equations (9)-(12) and (22) with $\hat{T}(\lambda,\kappa,\sigma) = \hat{L}(\lambda) = \hat{\mu}(\sigma) = \hat{q}(\kappa) = 1$ and $\hat{s}_{nn}(\kappa) = s_{nn,t}(\kappa)^{-1}$.*

This proposition follows trivially from the fact that changes in trade costs that move country n to autarky cause absorption shares for each κ to rise from $s_{nn,t}(\kappa)$ to 1, so that $\hat{s}_{nn}(\kappa) = s_{nn,t}(\kappa)^{-1}$. The effect on relative wages of moving to autarky depends on $\eta(\kappa)$ and $s_{nn,t}(\kappa)$. Intuitively, a high value of $s_{nn,t}(\kappa)$ implies a small effect of moving to autarky on the price of κ, since country n is not importing a large share of its investment in κ. A lower elasticity of substitution, $\eta(\kappa)$, so that the domestic investment good κ is a poor substitute for the imported variety, magnifies the impact of a given change in $s_{nn,t}(\kappa)$.

Whereas Proposition 3 provides a simple approach to quantify the impact on relative wages of moving country n to autarky, it does not directly shed light on the impact of international trade on inequality between two time periods t_0 and t_1. Our second counterfactual does. It answers the following question: What are the differential effects of changes in primitives (i.e. worldwide technologies, endowments, and trade costs) between periods t_0 and t_1 on wages in country n, relative to the effects of the same changes in primitives if country n were a closed economy? Answering this question seems difficult, because our estimation procedure does not recover all changes in country n's primitives (e.g. trade costs, foreign technologies, or endowments). Nevertheless, we can apply Proposition 3 to answer this question, as described in the following corollary.[28]

[28] To understand this result, define $w_n(\lambda;\Phi_t,\Phi^*{}_t,d_t)$ to be the average wage of worker type λ in country n given that country n parameters are Φ_t, parameters in the rest of the world are Φ^*_t, and the

Year	HSG/HSD	SMC/HSD	CLG/HSD	GTC/HSD	Skill premium
84	-0.000	-0.000	-0.001	-0.001	-0.001
03	-0.010	-0.019	-0.026	-0.027	-0.015

Table 9: The impact on log relative wages of moving to autarky in 1984 and 2003

Corollary 1. *The differential effects of changes in primitives between periods t_0 and t_1 on relative wages in country n, relative to the effects of the same changes in primitives if country n were a closed economy, are given by* $\frac{\hat{w}^A_{n,t_0}(\lambda) \big/ \hat{w}^A_{n,t_0}(\lambda_1)}{\hat{w}^A_{n,t_1}(\lambda) \big/ \hat{w}^A_{n,t_1}(\lambda)}$.

According to Corollary 1, we can quantify the effects on wages between periods t_0 and t_1 of international trade in country n following the same procedure described above, using only observed domestic absorption shares at time t_0 and t_1, rather than (unobserved) changes in primitives.

Results. Given our previous estimation, to conduct our counterfactuals we need only to assign values to $\eta(\kappa)$ and $s_{nn,t}(\kappa)$ for the United States in 1984 and 2003. We impose $\eta(\kappa_1) = \eta(\kappa_2)$ and set $\eta(\kappa) - 1 = 4.5$ to match a trade elasticity of 4.5 estimated in the equipment sector by Parro (2013). We calculate $s_{nn,t}(\kappa)$ for the U.S. as $s_{nn,t}(\kappa) = \frac{Production_{n,t}(\kappa) - Export_{n,t}(\kappa)}{Production_{n,t}(\kappa) - Export_{n,t}(\kappa) + Import_{n,t}(\kappa)}$ obtaining Production, Export, and Import data for $\kappa = \kappa_1, \kappa_2$ using the OECD's STAN STructural ANalysis Database (STAN), equating κ_1 in the model to industry 30 (Office, Accounting, and Computing Machinery) and κ_2 in the model to industries 29-33 less 30 (Machinery and Equipment less Office, Accounting, and computering Machinery) and 34-35 (Transport Equipment). We obtain similar domestic aborption shares in κ_1, $s_{nn,84}(\kappa_1) = 0.796$, and κ_2, $s_{nn,84}(\kappa_2) = 0.830$, in 1984. Whereas both domestic absorption shares fall between 1984 and 2003, the reduction in $s_{nn}(\kappa_1)$ is much larger. In 2003 we obtain $s_{nn,03}(\kappa_1) = 0.256$ and $s_{nn,03}(\kappa_2) = 0.650$.

Table 9 reports the effect on log relative wages of moving to autarky in 1984 and 2003 in the sectors and occupations models. The impact in 1984 is small, roughly 0.1 percentage points, regardless of aggregation (the skill premium vs. five education groups). The impact in 2003, while an order of magnitude larger than in 1984, is still small. Moving from 2003 to autarky generates a 1.5 percentage point reduction in the skill premium and a 2.7 percentage point reduction in the relative wage of the most educated to the least

full matrix of world trade costs are d_t. Define $d^A_{n,t}$ to be an alternative matrix of world trade costs in which country n's trade costs are infinite ($d_{in,t} = \infty$ for all $i \neq n$). We are interested in calculating $\left[\frac{w_n(\lambda;\Phi_{t'},\Phi^*_{t'},d_{t'})}{w_n(\lambda;\Phi_t,\Phi^*_t,d_t)}\right]\left[\frac{w_n\left(\lambda;\Phi_{t'},\Phi^*_{t'},d^A_{n,t'}\right)}{w_n\left(\lambda;\Phi_t,\Phi^*_t,d^A_{n,t}\right)}\right]^{-1}$. The result in Corollary 1 follows directly from Proposition 3 because $\frac{w_n(\lambda;\Phi_t,\Phi^*_t,d_t)}{w_n\left(\lambda;\Phi_t,\Phi^*_t,d^A_{n,t}\right)} = \left[\hat{w}^A_{n,t}(\lambda)\right]^{-1}$ for any time period.

educated group.

How important was trade in generating relative wage changes between 1984 and 2003? To answer this question, Corollary 1 states that we can simply difference the 2003 and 1984 results presented in Table 9. If the U.S. were in autarky between 1984 and 2003 but otherwise experienced the same changes in primitives, the U.S. skill premium would have risen by 1.4 percentage points less than it did over this time period. While this is a substantial difference, it accounts for only a small share of the total impact of changes in capital productivity—displayed in Table 4—and even less of the observed change in the U.S. skill premium over this time period.

The impact of trade on the skill premium is relatively small in the US, which has a comparative advantage in capital equipment. For the same estimated parameters, the impact of trade on the skill premium would be much larger in countries that rely heavily on imports for their capital equipment (see Burstein et al. (2013)).

9 Conclusions

To be added

References

Acemoglu, Daron, "Technical Change, Inequality, and the Labor Market," *Journal of Economic Literature*, March 2002, *40* (1), 7–72.

_ **and David Autor**, *Skills, Tasks and Technologies: Implications for Employment and Earnings*, Vol. 4 of *Handbook of Labor Economics*, Elsevier,

Arkolakis, Costas, Arnaud Costinot, and Andres Rodriguez-Clare, "New Trade Models, Same Old Gains?," *American Economic Review*, February 2012, *102* (1), 94–130.

Autor, David H., Frank Levy, and Richard J. Murnane, "The Skill Content Of Recent Technological Change: An Empirical Exploration," *The Quarterly Journal of Economics*, November 2003, *118* (4), 1279–1333.

_ **, Lawrence F. Katz, and Alan B. Krueger**, "Computing Inequality: Have Computers Changed The Labor Market?," *The Quarterly Journal of Economics*, November 1998, *113* (4), 1169–1213.

— , — , **and Melissa S. Kearney**, "Trends in U.S. Wage Inequality: Revising the Revisionists," *The Review of Economics and Statistics*, May 2008, *90* (2), 300–323.

Burstein, Ariel, Javier Cravino, and Jonathan Vogel, "Importing Skill-Biased Technology," *American Economic Journal: Macroeconomics*, April 2013, *5* (2), 32–71.

Caselli, Francesco and Daniel J. Wilson, "Importing technology," *Journal of Monetary Economics*, January 2004, *51* (1), 1–32.

Costinot, Arnaud and Jonathan Vogel, "Matching and Inequality in the World Economy," *Journal of Political Economy*, 08 2010, *118* (4), 747–786.

DiNardo, John E and Jorn-Steffen Pischke, "The Returns to Computer Use Revisited: Have Pencils Changed the Wage Structure Too?," *The Quarterly Journal of Economics*, February 1997, *112* (1), 291–303.

Eaton, Jonathan and Samuel Kortum, "Technology, Geography, and Trade," *Econometrica*, September 2002, *70* (5), 1741–1779.

Feenstra, Robert C. and Gordon H. Hanson, "The Impact Of Outsourcing And High-Technology Capital On Wages: Estimates For The United States, 1979-1990," *The Quarterly Journal of Economics*, August 1999, *114* (3), 907–940.

Goos, Maarten, Alan Manning, and Anna Salomons, "Job Polarization in Europe," *American Economic Review*, May 2009, *99* (2), 58–63.

Gordon, Robert J., *The Measurement of Durable Goods Prices*, National Bureau of Economic Research Research Monograph Series, 1990.

Greenwood, Jeremy, Zvi Hercowitz, and Per Krusell, "Long-Run Implications of Investment-Specific Technological Change," *American Economic Review*, June 1997, *87* (3), 342–62.

Grossman, Gene, Elhanan Helpman, and Philipp Kircher, "Matching and Sorting in a Global Economy," 2013. Working Paper Princeton University.

Grossman, Gene M. and Esteban Rossi-Hansberg, "Trading Tasks: A Simple Theory of Offshoring," *American Economic Review*, December 2008, *98* (5), 1978–97.

Helpman, Elhanan, Oleg Itskhoki, Marc-Andreas Muendler, and Stephen J. Redding, "Trade and Inequality: From Theory to Estimation," NBER Working Papers 17991, National Bureau of Economic Research, Inc April 2012.

Hornstein, Andreas, Per Krusell, and Giovanni L. Violante, "Frictional Wage Dispersion in Search Models: A Quantitative Assessment," *American Economic Review*, December 2011, *101* (7), 2873–98.

Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones, and Peter J. Klenow, "The Allocation of Talent and U.S. Economic Growth," NBER Working Papers 18693, National Bureau of Economic Research, Inc January 2013.

Huggett, Mark, Gustavo Ventura, and Amir Yaron, "Sources of Lifetime Inequality," *American Economic Review*, December 2011, *101* (7), 2923–54.

Kambourov, Gueorgui and Iourii Manovskii, "Occupational Mobility and Wage Inequality," *Review of Economic Studies*, 2009, *76* (2), 731–759.

Karabarbounis, Loukas and Brent Neiman, "The Global Decline of the Labor Share," NBER Working Papers 19136, National Bureau of Economic Research, Inc June 2013.

Krueger, Alan B, "How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984-1989," *The Quarterly Journal of Economics*, February 1993, *108* (1), 33–60.

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

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

McFadden, Daniel, *Frontiers of Econometrics*, New York, NY: Academic Press,

Parro, Fernando, "Capital-Skill Complementarity and the Skill Premium in a Quantitative Model of Trade," *American Economic Journal: Macroeconomics*, April 2013, *5* (2), 72–117.

Restuccia, Diego and Guillaume Vandenbroucke, "The Evolution of Education: A Macroeconomic Analysis," Working Papers tecipa-339, University of Toronto, Department of Economics October 2008.

Rothe, Christoph, "Decomposing the Composition Effect," IZA Discussion Papers 6397, Institute for the Study of Labor (IZA) February 2012.

_ , "Partial Distributional Policy Effects," *Econometrica*, 09 2012, *80* (5), 2269–2301.

A Appendix...

To be added

A.1 Proof of Proposition 2

Here we show how to conduct the decomposition between $t = t_0, t_1$ given (i) ρ, θ, and α; (ii) $L_t(\lambda)/L_t(\lambda_1)$; (iii) $\psi\tau(\lambda,\kappa,\sigma)$ for an arbitrary constant ψ; and (iv) $\tau_{\lambda t}(\lambda)/\tau_{\lambda t}(\lambda_1)$, $\tau_{\sigma t}(\sigma)/\tau_{\sigma t}(\sigma_1)$, and $\tau_{\kappa t}(\kappa)/\tau_{\kappa t}(\kappa_1)$.

Given $(i)-(iv)$ we construct $\pi_{t_0}(\lambda,\kappa,\sigma)$ and $w_{t_0}(\lambda)L_{t_0}(\lambda)/w_{t_0}(\lambda_1)L_{t_0}(\lambda_1)$ using equations (4), (5), and (14) as

$$\pi_{t_0}(\lambda,\kappa,\sigma) = \frac{\psi\tau(\lambda,\kappa,\sigma)\frac{\tau_{\lambda t_0}(\lambda)}{\tau_{\lambda t_0}(\lambda_1)}\frac{\tau_{\kappa t_0}(\kappa)}{\tau_{\kappa t_0}(\kappa_1)}\frac{\tau_{\sigma t_0}(\sigma)}{\tau_{\sigma t_0}(\sigma_1)}}{\sum_{\kappa',\sigma'}\psi\tau(\lambda,\kappa',\sigma')\frac{\tau_{\lambda t_0}(\lambda')}{\tau_{\lambda t_0}(\lambda_1)}\frac{\tau_{\kappa t_0}(\kappa')}{\tau_{\kappa t_0}(\kappa_1)}\frac{\tau_{\sigma t_0}(\sigma')}{\tau_{\sigma t_0}(\sigma_1)}}$$

and

$$\frac{w_{t_0}(\lambda)L_{t_0}(\lambda)}{w_{t_0}(\lambda_1)L_{t_0}(\lambda_1)} = \left(\frac{\sum_{\kappa,\sigma}\psi\tau(\lambda,\kappa,\sigma)\frac{\tau_{\lambda t_0}(\lambda)}{\tau_{\lambda t_0}(\lambda_1)}\frac{\tau_{\kappa t_0}(\kappa)}{\tau_{\kappa t_0}(\kappa_1)}\frac{\tau_{\sigma t_0}(\sigma)}{\tau_{\sigma t_0}(\sigma_1)}}{\sum_{\kappa',\sigma'}\psi\tau(\lambda_1,\kappa',\sigma')}\right)^{1/\theta}$$

Given (ii) we construct $\frac{\hat{L}(\lambda)}{\hat{L}(\lambda_1)} = \frac{L_{t_1}(\lambda)}{L_{t_0}(\lambda)}\frac{L_{t_0}(\lambda_1)}{L_{t_1}(\lambda_1)}$, and given (iv) we construct $\frac{\hat{\tau}_x(x)}{\hat{\tau}_x(x_1)} = \frac{\tau_{xt_1}(x)}{\tau_{xt_0}(x)}\frac{\tau_{xt_0}(x_1)}{\tau_{xt_1}(x_1)}$ for $x = \lambda,\kappa,\sigma$.

Given $(i)-(iv)$, we have

$$\frac{\zeta_{t_0}(\sigma)}{\zeta_{t_0}(\sigma_1)} = \frac{\sum_{\lambda,\kappa}\frac{w_{t_0}(\lambda)L_{t_0}(\lambda)}{w_{t_0}(\lambda_1)L_{t_0}(\lambda_1)}\pi_{t_0}(\lambda,\kappa,\sigma)}{\sum_{\lambda',\kappa'}\frac{w_{t_0}(\lambda')L_{t_0}(\lambda')}{w_{t_0}(\lambda_1)L_{t_0}(\lambda_1)}\pi_{t_0}(\lambda',\kappa',\sigma_1)}$$

so that

$$\frac{\hat{\zeta}(\sigma)}{\hat{\zeta}(\sigma_1)} = \frac{\zeta_{t_0}(\sigma_1)}{\zeta_{t_0}(\sigma)}\frac{\sum_{\lambda,\kappa}\frac{w_{t_0}(\lambda)L_{t_0}(\lambda)}{w_{t_0}(\lambda_1)L_{t_0}(\lambda_1)}\pi_{t_0}(\lambda,\kappa,\sigma)\frac{\hat{w}(\lambda)\hat{L}(\lambda)}{\hat{w}(\lambda_1)\hat{L}(\lambda_1)}\hat{\pi}(\lambda,\kappa,\sigma)}{\sum_{\lambda',\kappa'}\frac{w_{t_0}(\lambda')L_{t_0}(\lambda')}{w_{t_0}(\lambda_1)L_{t_0}(\lambda_1)}\pi_{t_0}(\lambda',\kappa',\sigma)\frac{\hat{w}(\lambda')\hat{L}(\lambda')}{\hat{w}(\lambda_1)\hat{L}(\lambda_1)}\hat{\pi}(\lambda',\kappa',\sigma_1)},$$

We conduct each exercise as follows, using the previously constructed variables.

Lemma 1. *Given changes in labor supplies, $\hat{L}(\lambda)/\hat{L}(\lambda_1)$, and values of $w_{t_0}(\lambda)L_{t_0}(\lambda)/(w_{t_0}(\lambda_1)L_{t_0}(\lambda_1))$ and $\pi_{t_0}(\lambda,\kappa,\sigma)$, changes in relative wages between t_0 and t_1 can be calculated using*

$$\frac{\hat{w}(\lambda)}{\hat{w}(\lambda_1)} = \left\{\frac{\sum_{\kappa,\sigma}\left(\hat{p}(\sigma)/\hat{p}(\sigma_1)\right)^{\frac{\theta}{1-\alpha}}\pi_{t_0}(\lambda,\kappa,\sigma)}{\sum_{\kappa',\sigma'}\left(\hat{p}(\sigma')/\hat{p}(\sigma_1)\right)^{\frac{\theta}{1-\alpha}}\pi_{t_0}(\lambda_1,\kappa',\sigma')}\right\}^{1/\theta}$$

$$\left(\frac{\hat{p}(\sigma)}{\hat{p}(\sigma_1)}\right)^{1-\rho} = \frac{\zeta_{t_0}(\sigma_1)}{\zeta_{t_0}(\sigma)}\frac{\sum_{\lambda,\kappa}\frac{w_{t_0}(\lambda)L_{t_0}(\lambda)}{w_{t_0}(\lambda_1)L_{t_0}(\lambda_1)}\pi_{t_0}(\lambda,\kappa,\sigma)\frac{\hat{w}(\lambda)\hat{L}(\lambda)}{\hat{w}(\lambda_1)\hat{L}(\lambda_1)}\hat{\pi}(\lambda,\kappa,\sigma)}{\sum_{\lambda',\kappa'}\frac{w_{t_0}(\lambda')L_{t_0}(\lambda')}{w_{t_0}(\lambda_1)L_{t_0}(\lambda_1)}\pi_{t_0}(\lambda',\kappa',\sigma)\frac{\hat{w}(\lambda')\hat{L}(\lambda')}{\hat{w}(\lambda_1)\hat{L}(\lambda_1)}\hat{\pi}(\lambda',\kappa',\sigma_1)}$$

$$\hat{\pi}(\lambda,\kappa,\sigma) = \frac{\left(\hat{p}(\sigma)/\hat{p}(\sigma_1)\right)^{\frac{\theta}{1-\alpha}}}{\sum_{\sigma',\kappa'}\left(\hat{p}(\sigma')/\hat{p}(\sigma_1)\right)^{\frac{\theta}{1-\alpha}}\pi_{t_0}(\lambda,\kappa',\sigma')}.$$

Lemma 1 follows directly.

Lemma 2. *Given changes in labor productivities, captured by $\hat{\tau}_\lambda(\lambda)$, and values of $w_{t_0}(\lambda) L_{t_0}(\lambda) / (w_{t_0}(\lambda_1) L_{t_0}(\lambda_1))$ and $\pi_{t_0}(\lambda,\kappa,\sigma)$, changes in relative wages between t_0 and t_1 can be calculated using*

$$\frac{\hat{w}(\lambda)}{\hat{w}(\lambda_1)} = \left\{ \frac{\hat{\tau}_\lambda(\lambda)}{\hat{\tau}_\lambda(\lambda_1)} \frac{\sum_{\kappa,\sigma} \left(\hat{p}(\sigma)/\hat{p}(\sigma_1)\right)^{\frac{\theta}{1-\alpha}} \pi_{t_0}(\lambda,\kappa,\sigma)}{\sum_{\kappa',\sigma'} \left(\hat{p}(\sigma')/\hat{p}(\sigma_1)\right)^{\frac{\theta}{1-\alpha}} \pi_{t_0}(\lambda_1,\kappa',\sigma')} \right\}^{1/\theta}$$

$$\left(\frac{\hat{p}(\sigma)}{\hat{p}(\sigma_1)}\right)^{1-\rho} = \frac{\zeta_{t_0}(\sigma_1)}{\zeta_{t_0}(\sigma)} \frac{\sum_{\lambda,\kappa} \frac{w_{t_0}(\lambda)L_{t_0}(\lambda)}{w_{t_0}(\lambda_1)L_{t_0}(\lambda_1)} \pi_{t_0}(\lambda,\kappa,\sigma) \frac{\hat{w}(\lambda)}{\hat{w}(\lambda_1)} \hat{\pi}(\lambda,\kappa,\sigma)}{\sum_{\lambda',\kappa'} \frac{w_{t_0}(\lambda')L_{t_0}(\lambda')}{w_{t_0}(\lambda_1)L_{t_0}(\lambda_1)} \pi_{t_0}(\lambda',\kappa',\sigma) \frac{\hat{w}(\lambda')}{\hat{w}(\lambda_1)} \hat{\pi}(\lambda',\kappa',\sigma_1)}$$

$$\hat{\pi}(\lambda,\kappa,\sigma) = \frac{\left(\hat{p}(\sigma)/\hat{p}(\sigma_1)\right)^{\frac{\theta}{1-\alpha}}}{\sum_{\sigma',\kappa'} \left(\hat{p}(\sigma')/\hat{p}(\sigma_1)\right)^{\frac{\theta}{1-\alpha}} \pi_{t_0}(\lambda,\kappa',\sigma')}.$$

Lemma 2 follows directly.

Lemma 3. *Given changes in capital productivities, captured by $\hat{\tau}_\kappa(\kappa)$, and values of $w_{t_0}(\lambda) L_{t_0}(\lambda) / (w_{t_0}(\lambda_1) L_{t_0}(\lambda_1))$ and $\pi_{t_0}(\lambda,\kappa,\sigma)$, changes in relative wages between t_0 and t_1 can be calculated using*

$$\frac{\hat{w}(\lambda)}{\hat{w}(\lambda_1)} = \left\{ \frac{\sum_{\kappa,\sigma} \left(\hat{p}(\sigma)/\hat{p}(\sigma_1)\right)^{\frac{\theta}{1-\alpha}} \left(\hat{\tau}_\kappa(\kappa)/\hat{\tau}_\kappa(\kappa_1)\right) \pi_{t_0}(\lambda,\kappa,\sigma)}{\sum_{\kappa',\sigma'} \left(\hat{p}(\sigma')/\hat{p}(\sigma_1)\right)^{\frac{\theta}{1-\alpha}} \left(\hat{\tau}_\kappa(\kappa')/\hat{\tau}_\kappa(\kappa_1)\right) \pi_{t_0}(\lambda_1,\kappa',\sigma')} \right\}^{1/\theta}$$

$$\left(\frac{\hat{p}(\sigma)}{\hat{p}(\sigma_1)}\right)^{1-\rho} = \frac{\zeta_{t_0}(\sigma_1)}{\zeta_{t_0}(\sigma)} \frac{\sum_{\lambda,\kappa} \frac{w_{t_0}(\lambda)L_{t_0}(\lambda)}{w_{t_0}(\lambda_1)L_{t_0}(\lambda_1)} \pi_{t_0}(\lambda,\kappa,\sigma) \frac{\hat{w}(\lambda)}{\hat{w}(\lambda_1)} \hat{\pi}(\lambda,\kappa,\sigma)}{\sum_{\lambda',\kappa'} \frac{w_{t_0}(\lambda)L_{t_0}(\lambda')}{w_{t_0}(\lambda_1)L_{t_0}(\lambda_1)} \pi_{t_0}(\lambda',\kappa',\sigma) \frac{\hat{w}(\lambda')}{\hat{w}(\lambda_1)} \hat{\pi}(\lambda',\kappa',\sigma_1)}$$

$$\hat{\pi}(\lambda,\kappa,\sigma) = \frac{\left(\hat{p}(\sigma)/\hat{p}(\sigma_1)\right)^{\frac{\theta}{1-\alpha}} \left(\hat{\tau}_\kappa(\kappa)/\hat{\tau}_\kappa(\kappa_1)\right)}{\sum_{\sigma',\kappa'} \left(\hat{p}(\sigma')/\hat{p}(\sigma_1)\right)^{\frac{\theta}{1-\alpha}} \left(\hat{\tau}_\kappa(\kappa')/\hat{\tau}_\kappa(\kappa_1)\right) \pi_{t_0}(\lambda,\kappa',\sigma')}.$$

Lemma 3 follows directly.

Lemma 4. *Given changes in task shifters, captured by $\hat{\tau}_\sigma(\sigma)$ and $\hat{\zeta}(\sigma)$, and values of $w_{t_0}(\lambda) L_{t_0}(\lambda) / (w_{t_0}(\lambda_1) L_{t_0}(\lambda_1))$ and $\pi_{t_0}(\lambda,\kappa,\sigma)$, changes in relative wages between t_0 and t_1 can be calculated using*

$$\frac{\hat{w}(\lambda)}{\hat{w}(\lambda_1)} = \left\{ \frac{\sum_{\kappa,\sigma} \left(\widetilde{p}(\sigma)/\widetilde{p}(\sigma_1)\right)^{\theta} \pi_{t_0}(\lambda,\kappa,\sigma)}{\sum_{\kappa',\sigma'} \left(\widetilde{p}(\sigma')/\widetilde{p}(\sigma_1)\right)^{\theta} \pi_{t_0}(\lambda_1,\kappa',\sigma')} \right\}^{1/\theta}$$

$$\frac{\hat{\zeta}_t(\sigma)}{\hat{\zeta}_t(\sigma_1)} \left(\frac{\hat{\tau}_\sigma(\sigma)}{\hat{\tau}_\sigma(\sigma_1)}\right)^{\frac{(\rho-1)(1-\alpha)}{\theta}} \left(\frac{\widetilde{p}(\sigma)}{\widetilde{p}(\sigma_1)}\right)^{(1-\rho)(1-\alpha)} = \frac{\sum_{\lambda,\kappa} \frac{w_{t_0}(\lambda)L_{t_0}(\lambda)}{w_{t_0}(\lambda_1)L_{t_0}(\lambda_1)} \pi_{t_0}(\lambda,\kappa,\sigma) \frac{\hat{w}(\lambda)}{\hat{w}(\lambda_1)} \hat{\pi}(\lambda,\kappa,\sigma)}{\sum_{\lambda',\kappa'} \frac{w_{t_0}(\lambda')L_{t_0}(\lambda')}{w_{t_0}(\lambda_1)L_{t_0}(\lambda_1)} \pi_{t_0}(\lambda',\kappa',\sigma) \frac{\hat{w}(\lambda')}{\hat{w}(\lambda_1)} \hat{\pi}(\lambda',\kappa',\sigma_1)}$$

$$\hat{\pi}(\lambda,\kappa,\sigma) = \frac{\left(\widetilde{p}(\sigma)/\widetilde{p}(\sigma_1)\right)^{\theta}}{\sum_{\sigma',\kappa'} \left(\widetilde{p}(\sigma')/\widetilde{p}(\sigma_1)\right)^{\theta} \pi_{t_0}(\lambda,\kappa',\sigma')},$$

where $\widetilde{p}(\sigma) \equiv \hat{p}(\sigma)^{\frac{1}{1-\alpha}} \hat{T}_{\sigma}(\sigma)$.

To prove Lemma 4, note that equation (10) and the definition of $\tau_{\sigma t}(\sigma)$ imply that between time t_0 and t_1, $\hat{T}_{\sigma}(\sigma)$ and $\hat{\mu}(\sigma)$ must satisfy the following condition

$$\frac{\hat{\mu}(\sigma)}{\hat{\mu}(\sigma_1)} = \left(\frac{\hat{T}_{\sigma}(\sigma_1)}{\hat{T}_{\sigma}(\sigma)} \left(\frac{\hat{\tau}_{\sigma}(\sigma)}{\hat{\tau}_{\sigma}(\sigma_1)} \right)^{\frac{1}{\theta}} \right)^{(\rho-1)(1-\alpha)} \frac{\hat{\zeta}(\sigma)}{\hat{\zeta}(\sigma_1)}, \tag{24}$$

Setting $\hat{L}(\lambda)$, $\hat{T}_{\lambda}(\lambda)$, $\hat{T}_{\kappa}(\kappa)$, and $\hat{q}(\kappa)$ all equal to one for each $(\lambda, \kappa, \sigma)$ in equations (9)-(12) and imposing condition (24), we obtain the system of equations in Lemma 4.

A.2 Data

To add

A.3 Estimation details

Here we explain in greater detail the estimation procedure in Section 5.2. As indicated in Proposition 2, once we have assigned values to the parameters ρ, θ and α (see Section 5.2.1), performing the decomposition described in Section 4 only requires identifying and estimating: (a) the comparative advantage parameter $\tau(\lambda, \kappa, \sigma)$ (up to an arbitrary scale parameter ψ); and, (b) the relative productivities $\tau_{xt}(x)/\tau_{xt}(x_0)$, for $x = \lambda, \kappa, \sigma$ and for every t.

In order to estimate the different components of the $\tau_t(\lambda, \kappa, \sigma)$ (see equation (14)), we will exclusively use data on $w_t(\lambda)$, and $\pi_t(\lambda, \kappa, \sigma)$, for every λ, κ, and σ. For each λ and period t, $w_t(\lambda)$ denotes the population average wage; however, our measure of $w_t(\lambda)$ is based on a sample average computed from a subset of a population of type λ. The same is true for $\pi_t(\lambda, \kappa, \sigma)$ for each triplet $(\lambda, \kappa, \sigma)$ and t, which denotes the population average of a dummy variable taking value 1 whenever a worker of type λ uses κ in σ.[29]

Given that we use sample averages to approximate population averages, in our estimation procedure, we allow for sampling error that generates differences between the unobserved population means, $w_t(\lambda)$ and $\pi_t(\lambda, \kappa, \sigma)$, and the observed sample averages, $w_t^*(\lambda)$ and $\pi_t^*(\lambda, \kappa, \sigma)$. We denote these errors as $\iota_{1t}(\lambda, \kappa, \sigma)$ and $\iota_2(\lambda)$:

$$\pi_t^*(\lambda, \kappa, \sigma) = \pi_t(\lambda, \kappa, \sigma)\iota_{1t}(\lambda, \kappa, \sigma), \qquad \forall\, (\lambda, \kappa, \sigma), \tag{25a}$$

$$w_t^*(\lambda) = w_t(\lambda)\iota_{2t}(\lambda), \qquad \forall\, \lambda. \tag{25b}$$

Given the Law of Large Numbers, for every t and $(\lambda, \kappa, \sigma)$, and for any real number $\xi > 0$, it holds

[29] Section 5.1 describes the sources of the data used to compute our measures of $w_t(\lambda)$, and $\pi_t(\lambda, \kappa, \sigma)$.

that

$$P\big[|\ln(\iota_{1t}(\lambda,\kappa,\sigma))| \geq \xi\big] \xrightarrow{N_t(\lambda)\to\infty} 0, \qquad \forall\ (\lambda,\kappa,\sigma), \tag{26a}$$

$$P\big[|\ln(\iota_{2t}(\lambda))| \geq \xi\big] \xrightarrow{N_t(\lambda)\to\infty} 0, \qquad \forall\ \lambda. \tag{26b}$$

Lemma 1. Given equations (14), (15), and (17) in the main text, and equation (25) and (26) in this Appendix, we can define an estimator $\delta(\lambda,\kappa,\sigma)$ such that, for every t, every (λ,κ,σ), and any real number $\xi > 0$,

$$P\left[\Big|\delta(\lambda,\kappa,\sigma) - \gamma^{\theta}\tau(\lambda,\kappa,\sigma)\Big| > \xi\right] \xrightarrow{N_t(\lambda)\to\infty} 0. \tag{27}$$

Proof. For every (λ,κ,σ), we define our estimator $\delta(\lambda,\kappa,\sigma)$ as

$$\ln(\delta(\lambda,\kappa,\sigma)) = \frac{1}{T}\sum_{t=1}^{T}\ln(w_t^*(\lambda)^{\theta}\pi_t^*(\lambda,\kappa,\sigma)). \tag{28}$$

From equation (25) in this Appendix, we can rewrite this expression as

$$\begin{aligned}\ln(\delta(\lambda,\kappa,\sigma)) &= \frac{1}{T}\sum_{t=1}^{T}\ln(w_t(\lambda)^{\theta}\pi_t(\lambda,\kappa,\sigma)\iota_{1t}(\lambda,\kappa,\sigma)\iota_{2t}(\lambda)) \\ &= \frac{1}{T}\sum_{t=1}^{T}\ln(w_t(\lambda)^{\theta}\pi_t(\lambda,\kappa,\sigma)) + \frac{1}{T}\sum_{t=1}^{T}\ln(\iota_{1t}(\lambda,\kappa,\sigma)) + \frac{1}{T}\sum_{t=1}^{T}\ln(\iota_{2t}(\lambda))\end{aligned}$$

From equations (16) and (17) in the main text, we can rewrite this expression in terms of $\tau(\lambda,\kappa,\sigma)$ as

$$\ln(\delta(\lambda,\kappa,\sigma)) = \theta\ln(\gamma) + \frac{1}{T}\sum_{t=1}^{T}\ln(\tau(\lambda,\kappa,\sigma)) + \frac{1}{T}\sum_{t=1}^{T}\ln(\iota_{1t}(\lambda,\kappa,\sigma)) + \frac{1}{T}\sum_{t=1}^{T}\ln(\iota_{2t}(\lambda))$$

Therefore, from equation (26) in this Appendix, we can conclude that

$$P\left[\left|\ln(\delta(\lambda,\kappa,\sigma)) - \Big(\theta\ln(\gamma) + \frac{1}{T}\sum_{t=1}^{T}\ln(\tau(\lambda,\kappa,\sigma))\Big)\right| > \xi\right] \xrightarrow{N_t(\lambda)\to\infty} 0,$$

or, equivalently,

$$P\left[\Big|\delta(\lambda,\kappa,\sigma) - \gamma^{\theta}\tau(\lambda,\kappa,\sigma)\Big| > \xi\right] \xrightarrow{N_t(\lambda)\to\infty} 0. \quad \blacksquare$$

Lemma 2. Given Lemma 1, equations (17), (18), (25a), and (26a), we can define an estimator $\delta_{\kappa t}(\kappa)$

such that, for every t, every κ, and any real number $\xi > 0$,

$$P\left[\left|\delta_{\kappa t}(\kappa) - \frac{\tau_{\kappa t}(\kappa)}{\tau_{\kappa t}(\kappa_1)}\right| > \xi\right] \xrightarrow{N_t(\lambda)\to\infty} 0. \tag{29}$$

Proof: For every κ, we define our estimator $\delta_{\kappa t}(\kappa)$ as:

$$\ln(\delta_{\kappa t}(\kappa)) = \frac{1}{n_\Lambda n_\Sigma} \sum_{\lambda,\sigma} \ln \frac{\pi_t^*(\lambda,\kappa,\sigma)}{\delta(\lambda,\kappa,\sigma)} \frac{\delta(\lambda,\kappa_1,\sigma)}{\pi_t^*(\lambda,\kappa_1,\sigma)} \tag{30}$$

From equation (25a), we can rewrite this expression as:

$$\ln(\delta_{\kappa t}(\kappa)) = \frac{1}{n_\Lambda n_\Sigma} \sum_{\lambda,\sigma} \ln \frac{\pi_t(\lambda,\kappa,\sigma)}{\delta(\lambda,\kappa,\sigma)} \frac{\delta(\lambda,\kappa_1,\sigma)}{\pi_t(\lambda,\kappa_1,\sigma)} + \frac{1}{n_\Lambda n_\Sigma} \sum_{\lambda,\sigma} \ln \frac{\iota_{1t}(\lambda,\kappa,\sigma)}{\iota_{1t}(\lambda,\kappa_1,\sigma)}.$$

From equation (17), we can rewrite this expression as:

$$\ln(\delta_{\kappa t}(\kappa)) = \frac{1}{n_\Lambda n_\Sigma} \sum_{\lambda,\sigma} \ln \frac{\tau_t(\lambda,\kappa,\sigma)}{\delta(\lambda,\kappa,\sigma)} \frac{\delta(\lambda,\kappa_1,\sigma)}{\tau_t(\lambda,\kappa_1,\sigma)} + \frac{1}{n_\Lambda n_\Sigma} \sum_{\lambda,\sigma} \ln \frac{\iota_{1t}(\lambda,\kappa,\sigma)}{\iota_{1t}(\lambda,\kappa_1,\sigma)}.$$

From Lemma 1 and equation (26a), we conclude that

$$P\left[\left|\ln(\delta_{\kappa t}(\kappa)) - \frac{1}{n_\Lambda n_\Sigma} \sum_{\lambda,\sigma} \ln \frac{\tau_t(\lambda,\kappa,\sigma)}{\tau(\lambda,\kappa,\sigma)} \frac{\tau(\lambda,\kappa_1,\sigma)}{\tau_t(\lambda,\kappa_1,\sigma)}\right| > \xi\right] \xrightarrow{N_t(\lambda,\kappa,\sigma)\to\infty} 0,$$

and, using equation (18),

$$P\left[\left|\ln(\delta_{\kappa t}(\kappa)) - \ln \frac{\tau_{\kappa t}(\kappa)}{\tau_{\kappa t}(\kappa_1)}\right| > \xi\right] \xrightarrow{N_t(\lambda)\to\infty} 0,$$

or, equivalently,

$$P\left[\left|\delta_{\kappa t}(\kappa) - \frac{\tau_{\kappa t}(\kappa)}{\tau_{\kappa t}(\kappa_1)}\right| > \xi\right] \xrightarrow{N_t(\lambda)\to\infty} 0. \quad \blacksquare$$

Lemma 3. Given Lemma 1, equations (17), (19), (25a), and (26a), we can define an estimator $\delta_{\sigma t}(\sigma)$

such that, for every t, every σ, and any real number $\xi > 0$,

$$P\left[\left|\delta_{\sigma t}(\sigma) - \frac{\tau_{\sigma t}(\sigma)}{\tau_{\sigma t}(\sigma_1)}\right| > \xi\right] \xrightarrow{N_t(\lambda)\to\infty} 0. \tag{31}$$

Proof: For every σ, we define our estimator $\delta_{\sigma t}(\sigma)$ as:

$$\ln(\delta_{\sigma t}(\sigma)) = \frac{1}{n_\Lambda n_K} \sum_{\lambda,\kappa} \ln \frac{\pi_t^*(\lambda,\kappa,\sigma)}{\delta(\lambda,\kappa,\sigma)} \frac{\delta(\lambda,\kappa,\sigma_1)}{\pi_t^*(\lambda,\kappa,\sigma_1)} \tag{32}$$

From equation (25a), we can rewrite this expression as:

$$\ln(\delta_{\sigma t}(\sigma)) = \frac{1}{n_\Lambda n_K} \sum_{\lambda,\kappa} \ln \frac{\pi_t(\lambda,\kappa,\sigma)}{\delta(\lambda,\kappa,\sigma)} \frac{\delta(\lambda,\kappa,\sigma_1)}{\pi_t(\lambda,\kappa,\sigma_1)} + \frac{1}{n_\Lambda n_K} \sum_{\lambda,k} \ln \frac{\iota_{1t}(\lambda,\kappa,\sigma)}{\iota_{1t}(\lambda,\kappa,\sigma_1)}.$$

From equation (17), we can rewrite this expression as:

$$\ln(\delta_{\sigma t}(\sigma)) = \frac{1}{n_\Lambda n_K} \sum_{\lambda,\kappa} \ln \frac{\tau_t(\lambda,\kappa,\sigma)}{\delta(\lambda,\kappa,\sigma)} \frac{\delta(\lambda,\kappa,\sigma_1)}{\tau_t(\lambda,\kappa,\sigma_1)} + \frac{1}{n_\Lambda n_K} \sum_{\lambda,\kappa} \ln \frac{\iota_{1t}(\lambda,\kappa,\sigma)}{\iota_{1t}(\lambda,\kappa,\sigma_1)}.$$

From Lemma 1 and equation (26a), we conclude that

$$P\left[\left|\ln(\delta_{\sigma t}(\sigma)) - \frac{1}{n_\Lambda n_K} \sum_{\lambda,k} \ln \frac{\tau_t(\lambda,\kappa,\sigma)}{\tau(\lambda,\kappa,\sigma)} \frac{\tau(\lambda,\kappa,\sigma_1)}{\tau_t(\lambda,\kappa,\sigma_1)}\right| > \xi\right] \xrightarrow{N_t(\lambda)\to\infty} 0,$$

and, using equation (19),

$$P\left[\left|\ln(\delta_{\sigma t}(\sigma)) - \ln \frac{\tau_{\sigma t}(\sigma)}{\tau_{\sigma t}(\sigma_1)}\right| > \xi\right] \xrightarrow{N_t(\lambda)\to\infty} 0,$$

or, equivalently,

$$P\left[\left|\delta_{\sigma t}(\sigma) - \frac{\tau_{\sigma t}(\sigma)}{\tau_{\sigma t}(\sigma_1)}\right| > \xi\right] \xrightarrow{N_t(\lambda)\to\infty} 0. \quad \blacksquare$$

Lemma 4 provides an alternative estimate of $\delta_{\lambda t}(\lambda)$.

Lemma 4. Given Lemma 1, equations (14), (17), (25a), and (26a), we can define an estimator $\delta(\lambda)$ such that, for every t, every λ, and any real number $\xi > 0$,

$$P\left[\left|\delta_{\lambda t}(\lambda) - \frac{\tau_{\lambda t}(\lambda)}{\tau_{\lambda t}(\lambda_1)}\right| > \xi\right] \xrightarrow{N_t(\lambda)\to\infty} 0. \tag{33}$$

Proof: For every λ, we define our estimator $\delta(\lambda)$ as:

$$\ln(\delta_{\lambda t}(\lambda)) = \frac{1}{n_\Sigma n_K} \sum_{\sigma,\kappa} \ln \frac{\pi_t^*(\lambda,\kappa,\sigma)}{\delta(\lambda,\kappa,\sigma)} \frac{\delta(\lambda_1,\kappa,\sigma)}{\pi_t^*(\lambda_1,\kappa,\sigma)} \tag{34}$$

From equation (25a), we can rewrite this expression as:

$$\ln(\delta_{\lambda t}(\lambda)) = \frac{1}{n_\Sigma n_K} \sum_{\sigma,\kappa} \ln \frac{\pi_t(\lambda,\kappa,\sigma)}{\delta(\lambda,\kappa,\sigma)} \frac{\delta(\lambda_1,\kappa,\sigma)}{\pi_t(\lambda_1,\kappa,\sigma)} + \frac{1}{n_\Lambda n_K} \sum_{\sigma,k} \ln \frac{\iota_{1t}(\lambda,\kappa,\sigma)}{\iota_{1t}(\lambda_1,\kappa,\sigma)}.$$

From equation (17), we can rewrite this expression as:

$$\ln(\delta_{\lambda t}(\lambda)) = \frac{1}{n_\Sigma n_K} \sum_{\sigma,\kappa} \ln \frac{\tau_t(\lambda,\kappa,\sigma)}{\delta(\lambda,\kappa,\sigma)} \frac{\delta(\lambda_1,\kappa,\sigma)}{\tau_t(\lambda_1,\kappa,\sigma)} + \frac{1}{n_\Lambda n_K} \sum_{\sigma,\kappa} \ln \frac{\iota_{1t}(\lambda,\kappa,\sigma)}{\iota_{1t}(\lambda_1,\kappa,\sigma)}.$$

From Lemma 1 and equation (26a), we conclude that

$$P\left[\left|\ln(\delta_{\lambda t}(\lambda)) - \frac{1}{n_\Sigma n_K} \sum_{\sigma,k} \ln \frac{\tau_t(\lambda,\kappa,\sigma)}{\tau(\lambda,\kappa,\sigma)} \frac{\tau(\lambda_1,\kappa,\sigma)}{\tau_t(\lambda_1,\kappa,\sigma)}\right| > \xi\right] \xrightarrow{N_t(\lambda)\to\infty} 0,$$

and, using equation (14) to obtain

$$\ln \frac{\tau_{\lambda t}(\lambda)}{\tau_{\lambda t}(\lambda_1)} = \frac{1}{n_\Sigma n_K} \sum_{\sigma,k} \ln \frac{\tau_t(\lambda,\kappa,\sigma)}{\tau(\lambda,\kappa,\sigma)} \frac{\tau(\lambda_1,\kappa,\sigma)}{\tau_t(\lambda_1,\kappa,\sigma)}$$

we, therefore, have

$$P\left[\left|\ln(\delta_{\lambda t}(\lambda)) - \ln \frac{\tau_{\lambda t}(\lambda)}{\tau_{\lambda t}(\lambda_1)}\right| > \xi\right] \xrightarrow{N_t(\lambda)\to\infty} 0,$$

or, equivalently,

$$P\left[\left|\delta_{\lambda t}(\lambda) - \frac{\tau_{\lambda t}(\lambda)}{\tau_{\lambda t}(\lambda_1)}\right| > \xi\right] \xrightarrow{N_t(\lambda)\to\infty} 0. \quad \blacksquare$$

Lemma 5. Given Lemmas 1, 2, and 3, equations (17), (20), (25b) and (26b), we can define an estimator $\delta^M(\lambda)$ such that, for every t, every λ, and any real number $\xi > 0$,

$$P\left[\left|\delta^M_{\lambda t}(\lambda) - \frac{\tau_{\lambda t}(\lambda)}{\tau_{\lambda t}(\lambda_1)}\right| > \xi\right] \xrightarrow{N_t(\lambda)\to\infty} 0. \tag{35}$$

Proof: For every λ, we define our estimator $\delta(\lambda)$ as:

$$\ln(\delta^M_{\lambda t}(\lambda)) = \theta \ln \frac{w^*_t(\lambda)}{w^*_t(\lambda_1)} - \ln \frac{\sum_{\kappa,\sigma} \delta_{\kappa t}(\kappa)\delta_{\sigma t}(\sigma)\delta(\lambda,\kappa,\sigma)}{\sum_{\kappa',\sigma'} \delta(\kappa')\delta(\sigma')\delta(\lambda_1,\kappa',\sigma')}. \tag{36}$$

From equation (25b), we can rewrite this expression as:

$$\ln(\delta^M_{\lambda t}(\lambda)) = \theta \ln \frac{w_t(\lambda)}{w_t(\lambda_1)} - \ln \frac{\sum_{\kappa,\sigma} \delta_{\kappa t}(\kappa)\delta_{\sigma t}(\sigma)\delta(\lambda,\kappa,\sigma)}{\sum_{\kappa',\sigma'} \delta(\kappa')\delta(\sigma')\delta(\lambda_1,\kappa',\sigma')} + \theta \ln \frac{\iota_{2t}(\lambda)}{\iota_{2t}(\lambda_1)}$$

From Lemmas 1, 2, and 3 and equation (26b), we conclude that

$$P\left[\left|\ln(\delta^M_{\lambda t}(\lambda)) - \left(\theta \ln \frac{w_t(\lambda)}{w_t(\lambda_1)} - \ln \frac{\sum_{\kappa,\sigma} \tau_{\kappa t}(\kappa)\tau_{\sigma t}(\sigma)\tau(\lambda,\kappa,\sigma)}{\sum_{\kappa',\sigma'} \tau_{\kappa t}(\kappa')\tau_{\sigma t}(\sigma')\tau(\lambda_1,\kappa',\sigma')}\right)\right| > \xi\right] \xrightarrow{N_t(\lambda)\to\infty} 0.$$

Corollary 1. Given Lemmas 1, 2, 3, 4, and 5, the estimators $\delta(\lambda,\kappa,\sigma)$, $\delta_{\lambda t}(\lambda)$, $\delta^M_{\lambda t}(\lambda)$, $\delta_{\kappa t}(\kappa)$, and $\delta_{\sigma t}(\sigma)$ such that, for every t, every (λ,κ,σ), and any real number $\xi > 0$,

$$P\left[\left\|\begin{pmatrix} \delta(\lambda,\kappa,\sigma) \\ \delta_{\lambda t}(\lambda) \\ \delta^M_{\lambda t}(\lambda) \\ \delta_{\kappa t}(\kappa) \\ \delta_{\sigma t}(\sigma) \end{pmatrix} - \begin{pmatrix} \gamma^{-\theta}\tau(\lambda,\kappa,\sigma) \\ \tau_{\lambda t}(\lambda)/\tau_{\lambda t}(\lambda_1) \\ \tau_{\lambda t}(\lambda)/\tau_{\lambda t}(\lambda_1) \\ \tau_{\kappa t}(\kappa)/\tau_{\kappa t}(\kappa_1) \\ \tau_{\sigma t}(\sigma)/\tau_{\sigma t}(\sigma_1) \end{pmatrix}\right\| > \xi\right] \xrightarrow{N_t(\lambda,\kappa,\sigma)\to\infty} 0. \tag{37}$$

Lemma 6. Given equation (28), it holds that $\delta(\lambda,\kappa,\sigma)/\delta(\lambda,\kappa,\sigma_1)$, and $\delta(\lambda,\kappa,\sigma)/\delta(\lambda,\kappa_1,\sigma)$, for all (λ,κ,σ), are independent of $w^*_t(\lambda)$, for every λ and every t.

Proof: Note that $\delta(\lambda,\kappa,\sigma)/\delta(\lambda,\kappa,\sigma_1)$ is independent of $w^*_t(\lambda)$, for every λ and every t, if and only if $\ln(\delta(\lambda,\kappa,\sigma)) - \ln(\delta(\lambda,\kappa,\sigma_1))$ is independent of $w^*_t(\lambda)$. From equation (28), note that

$$\begin{aligned}
\ln(\delta(\lambda,\kappa,\sigma)) - \ln(\delta(\lambda,\kappa,\sigma_1)) &= \frac{1}{T}\sum_{t=1}^{T} \ln(w^*_t(\lambda)^\theta \pi^*_t(\lambda,\kappa,\sigma)) - \frac{1}{T}\sum_{t=1}^{T} \ln(w^*_t(\lambda)^\theta \pi^*_t(\lambda,\kappa,\sigma_1)) \\
&= \frac{1}{T}\sum_{t=1}^{T} \ln(\pi^*_t(\lambda,\kappa,\sigma)) - \frac{1}{T}\sum_{t=1}^{T} \ln(\pi^*_t(\lambda,\kappa,\sigma_1))
\end{aligned}$$

and

$$\begin{aligned}
\ln(\delta(\lambda,\kappa,\sigma)) - \ln(\delta(\lambda,\kappa_1,\sigma)) &= \frac{1}{T}\sum_{t=1}^{T} \ln(w^*_t(\lambda)^\theta \pi^*_t(\lambda,\kappa,\sigma)) - \frac{1}{T}\sum_{t=1}^{T} \ln(w^*_t(\lambda)^\theta \pi^*_t(\lambda,\kappa_1,\sigma)) \\
&= \frac{1}{T}\sum_{t=1}^{T} \ln(\pi^*_t(\lambda,\kappa,\sigma)) - \frac{1}{T}\sum_{t=1}^{T} \ln(\pi^*_t(\lambda,\kappa_1,\sigma))
\end{aligned}$$

Corollary 2. Given equations (30) and (32), and Lemma 6, the estimators $\delta_{\kappa t}(\kappa)$, and $\delta_{\sigma t}(\sigma)$, for every t, every λ and every κ, are independent of the wage data and, therefore, also independent of the measurement error in wages, $\iota_{2t}(\lambda)$, for every λ and every t.

A.4 Comparative advantage in the data

Following Acemoglu and Autor (2011), we merge job task requirements from O*NET to their corresponding Census occupation classifications. We hold σ characteristics fixed over time.

We are interested in task characteristics to the extent that they shape worker and equipment comparative advantage across tasks. Hence, for our purposes the cleanest approach is to use directly a given number O*NET Work Activity and Work Context Importance scales, rather than aggregate these up, as in Acemoglu and Autor (2011), to form composite measures. We use the following 7 O*NET scales (with the O*NET code in paratheses): (i) Analyzing data/information (4.A.2.a.4); (ii) Thinking creatively (4.A.2.b.2); (iii) Guiding, directing, and motivating subordinates (4.A.4.b.4); (iv) Importance of repeating the same tasks (4.C.3.b.7); (v) Pace determined by speed of equipment (4.C.3.d.3); (vi) Manual dexterity (1.A.2.a.2); and (vii) Social Perceptiveness (2.B.1.a).

We normalize $\beta_{\lambda i}(\kappa_2) = \beta_{\sigma j}(\kappa_2) = 1$ for all i and j and estimate $\beta_{\lambda i}(\kappa)$ using variation in the share of worker types within each task using κ_2 relative to κ_1. According to equation (21), we have

$$\frac{1}{n_\Sigma}\sum_\sigma \log\left(\frac{\tau(\lambda,\kappa_1,\sigma)}{\tau(\lambda,\kappa_2,\sigma)}\Big/\frac{\tau(\lambda',\kappa_1,\sigma)}{\tau(\lambda',\kappa_2,\sigma)}\right) = \sum_{i=1}^{n_\lambda}\beta_{\lambda i}(\kappa_1)\left(X_i(\lambda) - X_i(\lambda')\right).$$

Whereas there are $n_\lambda = 4$ parameters and $n_\Lambda - 1 = 29$ observations for a given λ', we estimate this equation stacking observations for every possible λ' and adjust standard errors accordingly. We estimate $\beta_{\sigma j}(\kappa)$ symmetrically using

$$\frac{1}{n_\Lambda}\sum_\lambda \log\left(\frac{\tau(\lambda,\kappa_1,\sigma)}{\tau(\lambda,\kappa_2,\sigma)}\Big/\frac{\tau(\lambda,\kappa_1,\sigma')}{\tau(\lambda,\kappa_2,\sigma')}\right) = \sum_{j=1}^{n_\sigma}\beta_{\sigma j}(\kappa_1)\left(X_j(\sigma) - X_j(\sigma')\right).$$

Again, whereas there are $n_\sigma = 7$ parameters and $n_\Sigma - 1 = 19$ observations for a given σ', we estimate this equation stacking observations for every possible σ' and adjust standard errors accordingly. Finally, we estimate β_{ij} using

$$\frac{1}{n_K}\sum_\kappa \log\left(\frac{\tau(\lambda,\kappa,\sigma)}{\tau(\lambda,\kappa,\sigma')}\Big/\frac{\tau(\lambda',\kappa,\sigma)}{\tau(\lambda',\kappa,\sigma')}\right) = \sum_{i=1}^{n_\lambda}\sum_{j=1}^{n_\sigma}\beta_{ij}\left(X_i(\lambda) - X_i(\lambda')\right)\left(X_j(\sigma) - X_j(\sigma')\right),$$

similarly.

B Sectors

To add

Table 10: Occupations, their characteristics, and task shifters (1984-2003)

Occupations	Task shifter	Occupation characteristics						
		Data	Create	Guide	Repeat	Pace	Dext.	Social
Executive, administrative, managerial	1	5.45	5.41	6.13	4.95	1.77	1.80	7.42
Management related	0.99	5.80	4.79	4.70	5.73	1.59	1.86	6.63
Professional specialty	1.21	5.33	5.81	5.00	4.74	1.80	2.38	7.56
Technicians and related support	1.03	5.34	5.11	4.20	5.96	2.38	3.15	6.12
Financial sales and related	1.00	4.78	4.88	5.48	4.95	1.69	2.50	7.27
Retail sales	0.85	3.80	4.28	3.66	5.04	2.12	2.68	6.95
Administrative support	1.14	4.22	4.21	3.70	6.40	2.11	2.49	6.54
Housekeeping, cleaning, laundry	0.67	2.38	2.31	3.09	4.32	3.07	3.09	5.00
Protective service	0.84	4.63	4.51	4.94	6.09	1.98	3.46	7.14
Food preparation and service	1.13	3.22	3.78	4.07	4.62	2.72	3.75	6.60
Health service	1.23	3.54	4.21	3.55	5.10	2.03	3.23	7.19
Building, grounds cleaning, maintenance	0.82	2.80	3.89	3.55	4.04	2.79	3.85	5.78
Personal appearance, misc. personal care and service, recreation and hospitality	1.05	3.60	5.58	4.08	4.98	1.79	3.86	7.51
Child care	0.91	2.89	5.52	4.12	3.58	1.37	2.85	7.76
Farm operators and managers, other agricultural and related, extractive	0.63	4.34	4.25	4.23	4.41	3.60	4.20	5.56
Mechanics and repairers	0.67	4.49	4.76	4.25	4.60	2.66	4.52	5.71
Construction trades	0.62	4.11	4.70	4.84	4.40	2.84	4.12	5.74
Precision production	0.79	4.32	4.89	5.14	5.18	4.29	3.68	5.95
Machine operators, assemblers, inspectors	0.48	4.19	4.20	3.85	5.05	4.49	4.08	5.04
Transportation and material moving	0.75	3.74	3.95	3.65	4.88	3.64	4.02	5.80

Task shifter reports the change in task shifters between 1984 and 2003, evaluated at $\rho = 1$, and relative to the "Executive, administrative, managerial" occupation; Data: Analyzing data/information; Create: Thinking creatively; Guide: Guiding, directing, and motivating subordinates; Repeat: Importance of repeating the same tasks; Pace: Pace determined by speed of equipment; Dext.: Manual dexterity; Social: Social Perceptiveness.

C Interactions...