Accounting for Changes in Between-Group Inequality*

Ariel Burstein, UCLA
Eduardo Morales, Princeton University
Jonathan Vogel, Columbia University

February 26, 2014

Abstract

We provide a framework with multiple worker types (e.g. gender, age, education) to decompose changes in aggregated and disaggregated measures of between-group inequality into changes in (i) the composition of the workforce across labor types, (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 equipment types, worker types and tasks, and equipment types and tasks. We parameterize the model to match observed changes in worker type allocations (across equipment types and tasks) and wages in the United States between 1984 and 2003. The combination of changes in the importance of tasks and computerization explains more than half of the rise in the skill premium and the rise in inequality across more disaggregated education types as well as almost half the rise in the relative wage of women.

*We thank Treb Allen, Costas Arkolakis, Davin Chor, Pablo Fajgelbaum, Gene Grossman, Robert E. Lucas, Stephen Redding, Robert Shimer, and Nancy Stokey for helpful comments and David Autor and David Dorn for making their code available. 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 gap—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 changes in workforce composition, including rising relative supplies of educated workers and women; shifts in relative demand for workers across tasks (occupations and/or sectors), perhaps driven by structural transformation, international trade, or offshoring; computerization, evident from a rise in the stock and a fall in the price of computers relative to other capital equipment and structures; 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 measures of 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 measures of 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]

[1] In our discrete choice approach, factor allocation at a point in time provides information about aggre-

Comparative advantage (CA) shapes the allocation of labor to tasks and capital both directly and indirectly. For instance, the fact that one worker type λ uses computers relatively more than another λ' can be generated by two distinct patterns of comparative advantage. First, if λ has a comparative advantage using computers, then workers of type λ will use computers relatively more within tasks. We find that this is the case for more educated workers. Second, if λ has 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. We find that this is the case for women. 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 type, task pairs, 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 composition"), 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 ("capital 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), acts like an increase in the relative productivity of *workers* who have a direct comparative advantage using computers and like an increase in the relative productivity of *tasks* in which computers have a direct comparative advantage. Hence, if λ has a direct comparative advantage using computers, then this shock will tend to raise the relative wage of λ workers. On the other hand, if λ has a comparative advantage in the tasks in which computers have a comparative advantage, then this shock may either increase or decrease the relative wage of λ workers. Specifically, if tasks are complements, then a decline in the cost of producing computers will shrink employment in the tasks in which computers have a comparative advantage and will reduce the relative wages of workers with a comparative advantage in the same tasks as computers.[3] Thus, our model is flexible enough so that computerization may increase the relative wage of workers who are relatively productive using computers

gate 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.

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

[3]If tasks are substitutes, the opposite occurs.

and may reduce the relative wage of workers employed in tasks in which computers are particularly productive (even though both types of workers use computers relatively more than other worker types at the aggregate level), as described by, e.g., Autor et al. (1998) and Autor et al. (2003).

We can infer changes in equipment productivity, (iii), given data on changes over time in the usage of equipment types within worker type, task pairs. It is important to condition on worker type, task pairs because the usage of a particular type of equipment might go up for three reasons: it becomes relatively more productive or its cost falls (this is the change we want to identify), a shift in workforce composition towards workers who have a comparative advantage using that type of equipment, or a shift in task composition towards tasks in which that type of equipment has a comparative advantage. Similarly, we can infer changes in task shifters, (ii), given data on changes over time in task employment within worker type, equipment type pairs. Note that we identify changes in (ii) and (iii) without using data on wage changes, thus avoiding the critique of Acemoglu (2002) regarding previous work evaluating 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 decompose changes in between-group inequality into our four components.

Implementing this methodology requires data over time on the allocation of worker types to equipment type, task pairs. 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. First, it imposes a narrow view of computerization. Second, it only provides information on the allocation of workers to one type of equipment, computers; therefore, we must infer usage of the second type of equipment (non-computer equipment) in the model.[4] Finally, we do not observe the share of each worker's time at work spent using computers.[5]

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

[5]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.

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 decline in the price of computers relative to all other equipment types and structures, which we do not directly use in our estimation procedure.[6]

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 66% of the sum of the forces pushing the skill premium upward (the sum of task shifters, capital productivity, and labor productivity). 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 34% of the sum of the forces pushing the skill premium upward. In other words, observable changes in the allocation of workers

[6] 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).

to tasks and computers explain the majority of the rise in the skill premium. We also show that the combination of changes in capital productivity and task shifters account for almost half the forces reducing the gender gap. 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 a more disaggregated 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 (generating between-group wage polarization), decreasing wages of intermediate wage groups relative to the lowest and highest wage groups for our 15 groups of men between 1989 and 2003.[7]

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 results of our decomposition. Intuitively, shutting down any source(s) of comparative advantage affects how we infer the remaining source(s); e.g., if we abstract from worker-task comparative advantage, then to explain the fact that women use computers more intensively than men at the aggregate level, we must infer that they have a direct comparative advantage using computers. This suggests that modeling all three sources of comparative advantage is important for decomposing changes in between-group inequality. We also demonstrate the robustness of our results to a number of deviations from our baseline specification and parameterization. For instance, we find very similar results allowing for more general aggregate changes over time in the economic environment, such as changes in comparative advantage between labor and tasks.

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

[7] Autor et al. (2008) and the following literature study wage polarization over the full income distribution, whereas we focus on between-group inequality and, therefore, between-group wage polarization.

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 one fifth 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).[8]

We view our contributions as threefold. First, we nest four of the central mechanisms shaping relative wages proposed in the literature—labor composition, 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 framework 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 (i.e. between-group inequality), 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 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 (which we refer to as either equipment or as capital, for short), indexed by $\kappa \in \{\kappa_1, ..., \kappa_{n_K}\}$.[9] 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 $\Omega(\lambda) \subseteq \Omega$.

[8] 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.

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

Workers and capital equipment are employed by firms to produce tasks, indexed by $\sigma \in \{\sigma_1, ..., \sigma_{n_\Sigma}\}$.[10]

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 a 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 $0 \leq dep(\kappa) \leq 1$ 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 $[T_t(\lambda, \kappa, \sigma) \varepsilon_t(\omega, \kappa, \sigma)]^{1-\alpha(\sigma)} k^{\alpha(\sigma)}$.[11] The output elasticity of equipment in task σ is $\alpha(\sigma) \in (0, 1)$. In our baseline specification, the systematic component of productivity affecting all workers 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[12] $\{T_{xt}(x)\} \geq 0$ may vary over time for each $x \in \{\lambda, \kappa, \sigma\}$—so that some worker types, capital types, and tasks may become more productive than others over time[13]—

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

[11] 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.

[12] In the remainder of the paper, we denote by $\{f_t(x)\} = (f_t(x_1), ..., f_t(x_n))$ where x can take values from x_1 to x_n.

[13] Results are unchanged if we assume that the changing labor-specific component, $\{T_{\lambda t}(\lambda)\}$, combines

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.[14] 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.[15]

Relation to alternative frameworks. Whereas our framework imposes strong restrictions on production functions at the level of (λ,κ,σ), 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 referred 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-

productivity and a labor-specific distortion that creates a wedge between the wage received by the worker and the wage paid by the producer.

[14] Assuming that the time-varying components of productivity are separable between labor, tasks, and equipment types allows us to perform a clean decomposition between these forces. In Section 7.2, we allow for more general changes in technology over time: in the interaction between labor- and task-specific productivity, $\{T_{\lambda\sigma t}(\lambda,\sigma)\}$, or between labor- and equipment-specific productivity, $\{T_{\lambda\kappa t}(\lambda,\kappa)\}$, or between equipment- and task-specific productivity $\{T_{\kappa\sigma t}(\kappa,\sigma)\}$. In these cases we can only decompose changes in relative wages into three components. The results under these extensions do not vary substantially from our baseline results.

[15] Our analysis is unchanged if we introduce other homogeneous inputs (such as capital structures) 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.

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

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

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 $p_t(\sigma)\, k^{\alpha(\sigma)} \left[T_t(\lambda,\kappa,\sigma)\varepsilon_t(\omega,\kappa,\sigma)\right]^{1-\alpha(\sigma)} - k r_t(\kappa)$. 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 $\gamma(\lambda) \equiv \Gamma\left(1 - \frac{1}{\theta(\lambda)}\right)$ and $\Gamma(\cdot)$ is the Gamma function.

3.2 General equilibrium

In any period, task markets must clear,

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

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

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

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

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

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

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

$$\hat{w}(\lambda) = \hat{T}_\lambda(\lambda) \left\{ \sum_{\kappa,\sigma} \left[\left(\hat{r}(\kappa)^{-\alpha(\sigma)} \hat{p}(\sigma) \right)^{\frac{1}{1-\alpha(\sigma)}} \hat{T}_\kappa(\kappa) \hat{T}_\sigma(\sigma) \right]^{\theta(\lambda)} \pi_{t_0}(\lambda, \kappa, \sigma) \right\}^{1/\theta(\lambda)}, \tag{9}$$

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 given in equation (6),

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

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, $\frac{q_{t+1}(C)}{q_t(C)} \Big/ \frac{q_{t+1}(\kappa)}{q_t(\kappa)}$ for all κ, 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 $\{\theta(\lambda)\}$ in periods $t = t_0, t_1$, equations (9)-(13) determine changes in wages for any changes in technology, $\{q\hat{(\kappa)}\}$, $\{\hat{\mu}(\sigma)\}$, and $\{\hat{T}_x(x)\}$ for each $x \in \{\lambda,\kappa,\sigma\}$, 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 labor-equipment comparative advantage 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 labor-task and equipment-task comparative advantage 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 equation (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 impact 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 equation (9), changes in wages depend linearly—for given prices and rental rates—on changes in worker-specific productivities, $\{T_{\lambda t}(\lambda)\}$, and are a CES combination of changes in rental rates and prices, $\{r_t(\kappa)\}$ and $\{p_t(\sigma)\}$, 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 acts like a rise the relative productivity of 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 $\rho < 1$, a fall in the price of computers acts like an increase in the relative productivity of tasks in which computers have a CA, which leads to a reduction in employment (hence, hurts workers) in the tasks in which computers have a CA. 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 (and even though we do not allow for direct substitution of computers for workers in the production of individual tasks), 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.[16] Moreover, when $\rho < 1$ our empirical results are consistent with the findings in

[16]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.

Autor et al. (2003).[17]

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 composition*, $\left\{\hat{L}\left(\lambda\right)\right\}$,

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

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

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

Specifically, the direct contribution of changes in labor composition (for example) can be calculated 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, as reported in Section 6, these interaction effects are very small.

According to Proposition 1, we can conduct this decomposition exercise if we have values of parameters ρ, $\{\alpha(\sigma)\}$, and $\{\theta(\lambda)\}$, values of labor payments and allocations in period t, and measures of $\left\{\hat{L}\left(\lambda\right)\right\}$, $\left\{\hat{T}_{\lambda}\left(\lambda\right)\right\}$, $\left\{\hat{T}_{\sigma}\left(\sigma\right)\right\}$, $\{\hat{\mu}(\sigma)\}$, $\left\{\hat{T}_{\kappa}\left(\kappa\right)\right\}$, 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, 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 $\left\{\hat{T}_{\sigma}\left(\sigma\right)\right\}$, $\{\hat{\mu}(\sigma)\}$, $\left\{\hat{T}_{\kappa}\left(\kappa\right)\right\}$, 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

[17] 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 perform tasks in which thinking creatively is important) have a comparative advantage using computers. Hence, educated workers benefit (relatively) from a fall in computer prices.

equations (2) and (3), we have

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

where

$$\tau_{\lambda t}(\lambda) \equiv \left[\frac{T_{\lambda t}(\lambda)}{\prod_{t'=1}^{T} T_{\lambda t'}(\lambda)}\right]^{\theta} \tag{14b}$$

$$\tau_{\kappa t}(\kappa) \equiv \left[\frac{T_{\kappa t}(\kappa)\, r_t(\kappa)^{\frac{-\alpha}{1-\alpha}}}{\prod_{t'=1}^{T} T_{\kappa t'}(\kappa)\, r_{t'}(\kappa)^{\frac{-\alpha}{1-\alpha}}}\right]^{\theta} \tag{14c}$$

$$\tau_{\sigma t}(\sigma) \equiv \left[\frac{T_{\sigma t}(\sigma)\, p_t(\sigma)^{\frac{1}{1-\alpha}}}{\prod_{t'=1}^{T} T_{\sigma t'}(\sigma)\, p_{t'}(\sigma)^{\frac{1}{1-\alpha}}}\right]^{\theta} \tag{14d}$$

and

$$\tau(\lambda,\kappa,\sigma) \equiv \left[\frac{(1-\alpha)\,T(\lambda,\kappa,\sigma)}{\alpha^{\alpha/(\alpha-1)}} \prod_{t'=1}^{T} T_{\lambda t'}(\lambda)\, T_{\kappa t'}(\kappa)\, r_{t'}(\kappa)^{\frac{-\alpha}{1-\alpha}}\, T_{\sigma t'}(\sigma)\, p_{t'}(\sigma)^{\frac{1}{1-\alpha}}\right]^{\theta}. \tag{14e}$$

We now describe how to conduct the decomposition described above for given values, in $t = t_0, t_1$, of (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) $\{\tau_{xt}(x)/\tau_{\lambda t}(x_1)\}$ for all $x \in \{\lambda,\kappa,\sigma\}$. 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 (14a). Given (ii) we construct $\{\hat{L}(\lambda)/\hat{L}(\lambda_1)\}$ and given (iv) we construct $\{\hat{\tau}_x(x)/\hat{\tau}_\lambda(x_1)\}$ for all $x \in \{\lambda,\kappa,\sigma\}$. Using these, we conduct each exercise as follows. Details are provided in the Appendix.

Labor composition. To quantify the direct impact of changes in labor composition, we set $\hat{T}_\lambda(\lambda)$, $\hat{T}_\kappa(\kappa)$, $\hat{T}_\sigma(\sigma)$, $\hat{\mu}(\sigma)$, and $\hat{q}(\kappa)$ all equal to one for each λ,κ,σ, impose $\frac{\hat{L}(\lambda)}{\hat{L}(\lambda_1)} = \frac{L_{t_1}(\lambda)/L_{t_0}(\lambda)}{L_{t_1}(\lambda_1)/L_{t_0}(\lambda_1)}$ for all λ, 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 $\hat{L}(\lambda)$, $\hat{T}_\kappa(\kappa)$, $\hat{T}_\sigma(\sigma)$, $\hat{\mu}(\sigma)$, and $\hat{q}(\kappa)$ all equal to one for each λ,κ,σ, impose $\hat{T}_\lambda(\lambda)/\hat{T}_\lambda(\lambda_1) = (\hat{\tau}_\lambda(\lambda)/\hat{\tau}_\lambda(\lambda_1))^{1/\theta}$ for all λ, and use equations (9)-(13) to solve for changes in relative wages, as summarized by Lemma 2 in the Appendix.

Capital productivity. While both $\{\hat{T}_\kappa(\kappa)\}$ and $\{\hat{q}(\kappa)\}$ shape capital productivity, we cannot separately recover $\{\hat{T}_\kappa(\kappa)\}$ and $\{\hat{q}(\kappa)\}$ from $\{\hat{\tau}_\kappa(\kappa)\}$ without additional data. However, changes in relative wages depend on $\{\hat{T}_\kappa(\kappa)\}$ and $\{\hat{q}(\kappa)\}$ only through $\frac{\hat{\tau}_\kappa(\kappa)}{\hat{\tau}_\kappa(\kappa_1)} =$

$\left(\frac{\hat{T}_\kappa(\kappa')}{\hat{T}_\kappa(\kappa_1)}\left(\frac{\hat{q}(\kappa)}{\hat{q}(\kappa_1)}\right)^{\frac{-\alpha}{1-\alpha}}\right)^\theta$ for all κ. Hence, to quantify the direct impact of changes in labor productivity, we set $\hat{L}(\lambda)$, $\hat{T}_\lambda(\lambda)$, $\hat{T}_\sigma(\sigma)$, and $\hat{\mu}(\sigma)$ all equal to one for each λ, κ, σ, impose $\frac{\hat{T}_\kappa(\kappa)}{\hat{T}_\kappa(\kappa_1)}\left(\frac{\hat{q}(\kappa)}{\hat{q}(\kappa_1)}\right)^{-\alpha/(1-\alpha)} = \left(\frac{\hat{\tau}_\kappa(\kappa)}{\hat{\tau}_\kappa(\kappa_1)}\right)^{1/\theta}$ for each κ, 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 cost of producing equipment, $\{\hat{q}(\kappa)\}$, from changes in the productivity of using equipment, $\{\hat{T}_\kappa(\kappa)\}$, our estimates of $\{\hat{\tau}_\kappa(\kappa)\}$ are sufficient to conduct this exercise.

Task shifters. As in the capital productivity exercise, we cannot separate $\{\hat{T}_\sigma(\sigma)\}$ and $\{\hat{\mu}(\sigma)\}$ from $\{\hat{\tau}_\sigma(\sigma)\}$ without additional data. Nevertheless, we can quantify the direct impact of changes in task shifters by setting $\hat{L}(\lambda)$, $\hat{T}_\lambda(\lambda)$, $\hat{T}_\kappa(\kappa)$, and $\hat{q}(\kappa)$ all equal to one for each λ, κ, σ, and using equations (9)-(13) to solve for changes in relative wages for given $\{\hat{\tau}_\sigma(\sigma)\}$ and for given changes in incomes accruing in each task, $\{\hat{\zeta}(\sigma)\}$, as summarized by Lemma 4 in the Appendix. Whereas changes in the relative importance of tasks in the production of the unique final good, $\{\hat{\mu}(\sigma)\}$, 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 $t = t_0, t_1$—$\{L_t(\lambda)/L_t(\lambda_1)\}$, $\{\psi\tau(\lambda,\kappa,\sigma)\}$ for an arbitrary constant $\psi \neq 0$, and $\{\tau_{xt}(x)/\tau_{xt}(x_1)\}$ for all $x \in \{\lambda,\kappa,\sigma\}$, 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 our baseline exercises, we map tasks in the model to occupations in the data (in Appendix E we conduct our decomposition mapping tasks in the model to sectors in the data). We aggregate up to twenty occupations.[18] See Table 11 in Appendix B for a list of these occupations.

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

[18]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 (λ,κ,σ) and t for which $\pi_t(\lambda,\kappa,\sigma) = 0$.

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 by dropping workers who are younger than 17 years old, do not report positive (paid) hours worked, or are self-employed.[19] Here we briery describe our use of these sources; we provide further details in Appendix B.

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 composition, $\{L_t(\lambda)\}$, is hours worked within each labor type λ, and our measure of the average (hourly) wage, $\{w_t(\lambda)\}$, is total labor income in the previous year divided by total hours worked within each λ. See Appendix B 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 labor types.[20] 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 composition 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 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. In this sense, ours is a narrow measure of computerization.

[19] Our results are robust to including the self-employed (and measuring income with the sum of labor, business, and farm income).

[20] 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.

		1984	1989	1993	1997	2003
All		26.7	38.5	48.2	51.9	56.9
Gender	Men	22.9	32.7	42.2	45.4	50.9
	Women	32.0	46.1	55.8	60.2	64.3
Age	16-30	26.2	35.8	43.0	46.3	48.8
	31-43	30.9	43.8	52.9	55.6	59.9
	44 +	22.4	35.0	47.5	52.8	59.9
Education	HSD	5.26	7.38	10.5	12.6	15.5
	HSG	21.1	30.4	35.8	37.7	40.7
	SMC	34.8	48.1	55.8	58.6	59.6
	CLG	45.0	61.6	72.8	78.1	84.1
	GTC	45.4	61.3	74.9	81.8	88.9

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

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 almost every education category and year, more educated workers 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: $\kappa_1 = \mathit{computers}$ and $\kappa_2 = \mathit{other\ equipment}$. In our baseline model in which a worker uses exactly one type of equipment, we construct $\pi_t(\lambda, \kappa, \sigma)$ for $\kappa = \mathit{computers}$ 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 $\pi_t(\lambda, \kappa, \sigma)$ as we do introduces two issues. First, at the individual level our computer-use variable takes only two values: zero or one. In practice, a worker may use a computer during some, but not all of her workday. Ideally, we would want to allocate each worker's hours across equipment using more disaggregated data. The 2006 German *Qualification and Working Conditions* survey circumvents this issue by asking "How much of your total work time do you spend on computers?" Using this German data, we constructed $\pi_t(\lambda, \kappa_1, \sigma)$ 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;[21] doing so, we found similar patterns of comparative advantage as in the U.S. data. Second, we are not using any information on the allocation of non-computer capital equipment. In the version of our model in which tasks are mapped to sectors in the data (see Appendix E) 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 E 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 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 decompositions. 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

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

determines the dispersion of idiosyncratic productivity draws. As discussed in Lagakos and Waugh (2013) and Hsieh et al. (2013), the dispersion of wages across workers within a labor group λ who have chosen the same allocation (in our case the pair κ, 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 and $\{\tau_{xt}(x)/\tau_{xt}(x_1)\}$ for all $x \in \{\lambda,\kappa,\sigma\}$ and all t (where t =1984, 1989, 1993, 1999, and 2003) using allocation and wage data. Specifically, we estimate 1,444 parameters: $n_\Lambda n_K n_\Sigma - 1 = 1,199$ time-invariant parameters $\{\tau(\lambda,\kappa,\sigma)\}$, $(n_\Lambda - 1) \times 5$ years $=$ 29×5 parameters $\{\tau_{\lambda t}(\lambda)/\tau_{\lambda t}(\lambda_1)\}$, $(n_K - 1) \times 5 = 5$ parameters $\{\tau_{\kappa t}(\kappa)/\tau_{\kappa t}(\kappa_1)\}$, and $(n_\Sigma - 1) \times 5 = 19 \times 5$ parameters $\{\tau_{\sigma t}(\sigma)/\tau_{\sigma t}(\sigma_1)\}$. We do so using 5,995 observations: $(n_\Lambda - 1) \times 5 = 29 \times 5$ relative wage observations and $(n_\Lambda \times (n_K n_\Sigma - 1)) \times 5 = 1,170 \times 5$ allocation observations. We provide details in Appendix C.

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.[22]

Our estimation involves three steps. In the first step we estimate the parameters that determine comparative advantage, $\{\psi\tau(\lambda,\kappa,\sigma)\}$. In the second step we estimate equipment productivity and task shifters for each year, $\{\tau_{\kappa t}(\kappa)/\tau_{\kappa t}(\kappa_1)\}$ and $\{\tau_{\sigma t}(\sigma)/\tau_{\sigma t}(\sigma_1)\}$.[23]

[22] We have between roughly 46,000 and 54,000 workers in the October Supplement. With 30 labor types, this implies an average of about 1,700 observations per-λ in each year.

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

In the final step we estimate labor productivity for each year, $\{\tau_{\lambda t}(\lambda) / \tau_{\lambda t}(\lambda_1)\}$, using the estimates from steps one and two.

Step 1: Comparative advantage. Equations (14a)-(14e) give 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 $x \in \{\lambda, \kappa, \sigma\}$, 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 C. 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 C 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 these concerns. Specifically, equation (14a) 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 C 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 (14a) 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, controlling for worker comparative advantage, equipment productivities, and task shifters. In Appendix C 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 $\{\tau_{\lambda t}(\lambda)/\tau_{\lambda t}(\lambda_1)\}$.[24]

Equations (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, there are 902 $\pi_t(\lambda,\kappa,\sigma)$ observations (out of 5850) that are zeros.[25] If $\pi_t(\lambda,\kappa,\sigma)=0$ for all t, then—consistent with the model—we set $\tau(\lambda,\kappa,\sigma)=0$. The sample averages in equations (18) and (19) (and analogously for equation (16) when not all $\pi_t(\lambda,\kappa,\sigma)=0$) are computed using data on only the positive $\pi_t(\lambda,\kappa,\sigma)$.[26]

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 changes in relative wages between periods t_0 and t_1 do not directly

[24] Appendix C also describes an alternative strategy for estimating $\tau_{\lambda t}(\lambda)/\tau_{\lambda t}(\lambda_1)$ without exactly matching relative wages. Using this alternative approach yields similar results in our decomposition.

[25] In particular, the fraction of observations that are equal to 0 in the data are roughly: 18% for 1984, 15% in 1989 and 1997, 16% in 1993, and 11% in 2003.

[26] 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—the number of zero values becomes minimal. 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.

affect the strength of the labor composition, capital productivity, or task shifters components of our decomposition. This is in contrast to the labor productivity component, since $\{\tau_{\lambda t}(\lambda)/\tau_{\lambda t}(\lambda_1)\}$ is chosen as a residual to match relative wages in every period.

The labor composition 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 $\{\tau_{\kappa t_1}(\kappa)/\tau_{\kappa t_1}(\kappa')\}$ is independent of wages in period t_1, as shown in Appendix C. 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 C, as well as the share of total labor income in each task in period t_1, but does not depend directly on wages in period t_1.

Since changes in wages are not an input in the algorithms to determine the direct contributions of these components, there is no *a priori* reason to believe that the labor composition, 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.

5.3 Estimation results

By design our estimation procedure exactly matches relative wages in every period. The correlation between the model's implied $\{\pi_t(\lambda,\kappa,\sigma)\}$ and the $\{\pi_t(\lambda,\kappa,\sigma)\}$ in the data is 0.97, 0.98, 0.99, 0.99, and 0.98 in 1984, 1989, 1993, 1997, and 2003 respectively. In the remainder of 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 restrictions on how these vary across (λ,κ,σ). Solely for the purpose of summarizing 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); $\beta = (\beta_{11}, ..., \beta_{n_\lambda n_\sigma})$ is a vector with $n_\lambda n_\sigma$ elements; and $\beta_\lambda(\kappa) = (\beta_{\lambda 1}(\kappa), ..., \beta_{\lambda n_\lambda}(\kappa))$ and $\beta_\sigma(\kappa) = (\beta_{\sigma 1}(\kappa), ..., \beta_{\sigma n_\sigma}(\kappa))$ 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) relative to κ' (e.g. non-computers) is relatively more productive in tasks characterized by a high value of characteristic j (e.g., the importance of repeating the same task).

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 (e.g. college graduates have 16 years of education). 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 B. We use 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.[27] Table 11 in Appendix B lists these scales for each of the twenty occupations.

In Appendix C 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 that refer to worker characteristic

[27] 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."

$X(\lambda)$	$\beta_\lambda(\kappa_1) - \beta_\lambda(\kappa_2)$	$X(\sigma)$	$\beta_\sigma(\kappa_1) - \beta_\sigma(\kappa_2)$	$\beta_{education \times j}$
Age	-0.003	Analyzing data/information	0.135	0.247[a]
Female	0.261[a]	Thinking creatively	0.536[a]	0.054
Education	0.317[a]	Guiding, directing, motivating	0.050	-0.125[a]
		Importance of repetition	0.556[a]	-0.100[a]
		Pace determined equipment	-0.425[a]	-0.034
		Manual dexterity	-0.663[a]	-0.096[a]
		Social Perceptiveness	-0.672[a]	0.081

Table 2: Comparative advantage

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

$i =$ education. This table highlights four 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 $\exp(0.317 \times 4/\theta) = 1.51$ times more productive with computers (relative to non-computer equipment) than the one with a high school degree (12 years of education). Second, whereas females have a comparative advantage using computers, this comparative advantage is weak, in the sense that a female is about $\exp(0.261/\theta) = 1.09$ times more productive than an otherwise identical male using computers relative to non-computer equipment. Third, computers are relatively productive in tasks in which repetition (as suggested by results in Autor and Dorn (2013)) and thinking creatively 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.

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 described in the introduction.

Table 3 also summarizes changes in worker productivity. We aggregate up from thirty

	84-89	89-93	93-97	97-03	**84-03**
Capital					
κ_1/κ_2	0.54	0.46	0.09	0.27	**1.36**
Education					
HSG/HSD	-0.01	0.00	-0.01	-0.01	**-0.01**
SMC/HSD	0.01	-0.04	0.00	-0.03	**-0.02**
CLG/HSD	0.07	0.08	0.02	0.08	**0.26**
GTC/HSD	0.09	0.35	0.04	0.13	**0.63**

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

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 occurring after 1989.

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, in general 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 $\{\hat{\zeta}(\sigma)/\hat{\zeta}(\sigma_1)\}$.[28] Regressing $\{\hat{\zeta}(\sigma)/\hat{\zeta}(\sigma_1)\}$ between 1984 and 2003—reported in Table 11 in Appendix D—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 women have a comparative advantage) grew. Hence, we find that task shifters rose in tasks in which educated workers and women have a comparative advantage relative to in tasks in which they have a comparative disadvantage.

[28] 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 comp.	Task shifters	Labor prod.	Capital prod.
84-89	0.068	-0.025	0.042	0.019	0.030
89-93	0.078	-0.021	0.027	0.045	0.025
93-97	0.016	-0.009	0.013	0.007	0.005
97-03	0.050	-0.028	0.036	0.030	0.013
84-03	**0.209**	**-0.084**	**0.119**	**0.099**	**0.067**

Table 4: Decomposing changes in the log skill premium

6 Results

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

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 composition, 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, the combination of changes in capital productivity and task shifters explain the majority of the rise in the skill premium; see the final row of Table 4. The capital productivity component alone accounts for roughly 32% of the rise in the skill premium ($0.32 \simeq 0.067/0.209$) and roughly 24% of the sum of the forces pushing the skill premium upwards ($0.24 \simeq 0.067/(0.119 + 0.099 + 0.067)$). Over sub-periods, changes in capital productivity are particularly important in generating changes in the skill premium over 1984-1989 and 1989-1993.[29] 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).[30]

[29] Because we adjust the weighting of labor groups within college educated and non-college educated workers for each set of years considered, the sum of the columns for the sub-periods need not equal the value over the full sample.

[30] DiNardo and Pischke (1997) critique Krueger (1993) by showing that pencils can explain wage premia

	Data	Labor comp.	Task shifters	Labor prod.	Capital prod.
HSG/HSD	0.040	-0.029	0.049	-0.010	0.036
SMC/HSD	0.081	-0.061	0.101	-0.025	0.071
CLG/HSD	0.207	-0.109	0.162	0.047	0.107
GTC/HSD	0.338	-0.137	0.203	0.152	0.111

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

The task shifter component accounts for roughly 57% of the rise in the skill premium and 42% of the sum of the forces pushing the skill premium upwards over the full sample. We obtain the result that task shifters have substantially increased the U.S. skill premium 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.

Whereas the relative importance of changes in capital productivity and task shifters depends on our assigned value of ρ, as we describe in Section 7, the conclusion that the combination of changes in capital productivity and task shifters explains the majority of the rise in the skill premium is robust. Perhaps surprisingly, of the mechanisms pushing the skill premium upwards over the full sample, the one mechanism that was estimated to match observed relative wages (and, therefore, changes in relative wages), labor productivity, only accounts for roughly 47% of the rise in the skill premium and roughly 35% 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 forces increasing the relative wage of more educated workers.

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.

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 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 "between-group wage polarization," a between-group version of 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 composition, 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.[31]

Gender wage gap. Between 1984 and 2003 the log change in the composition adjusted gender wage gap (the average wage of males relative to females) in the data was -0.137. According to our decomposition, the rise in female labor supply increased the gender wage gap by 0.028 log points, task shifters decreased it by -0.072 log points, labor productivity decreased it by -0.085 log points, and capital productivity decreased it by -0.001 log points. These numbers highlight two important results. First, the combination of changes in capital productivity and task shifters explain almost half of the decline in the gender wage gap between 1984 and 2003; together they account for roughly 46% of the forces decreasing the gender wage gap. Second, capital productivity has almost no effect on the gender wage gap in spite of the fact that women are substantially more likely to use a computer at work than men, as shown in Table 1. As we show in Section 7.1, 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

[31] 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.

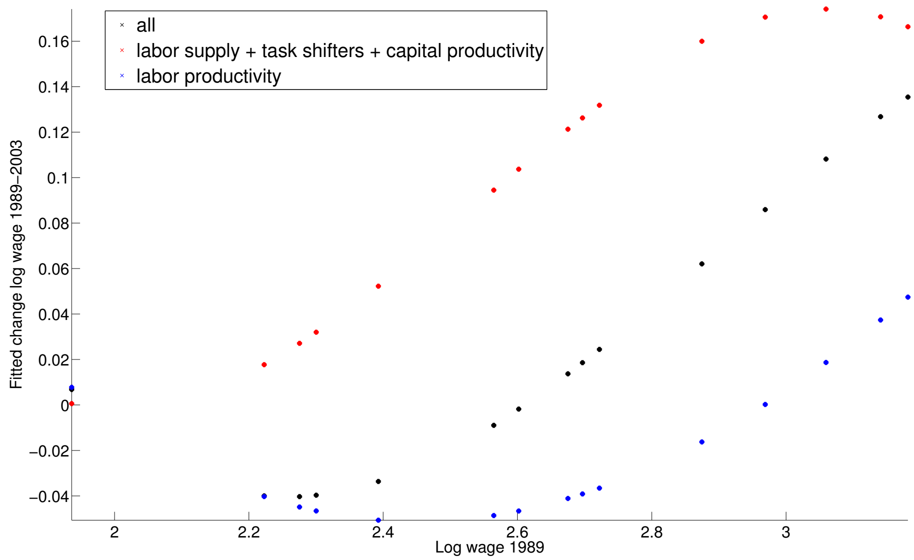

Figure 1: Cubic fit of the log change in average hourly wages between 1989-2003 (relative to the lowest wage group in 1989) plotted against log average hourly wages in 1989.

level, because 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 three types of sensitivity exercises. First, we illustrate the importance of all three types of comparative advantage by turning some of them off. Second, we allow for changes in comparative advantage over time. Finally, we perform sensitivity to different values of ρ and θ.

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 $\tau(\lambda, \kappa_1, \sigma) = \tau(\lambda, \kappa_2, \sigma)$ for all (λ, σ). This is equivalent, in terms of the model's implications for changes in rel-

	Skill premium				Gender gap			
	Labor comp.	Task shifters	Labor prod.	Capital prod.	Labor comp.	Task shifters	Labor prod.	Capital prod.
Baseline	-0.084	0.119	0.099	0.067	0.028	-0.072	-0.085	-0.001
Only labor-equipment CA	0	0	0.042	0.167	0	0	-0.079	-0.058
Only labor-task CA	-0.090	0.134	0.156	0	0.029	-0.059	-0.101	0

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

ative 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 all (λ, κ) and all $i = 2, ..., n_\Sigma$. This is equivalent—again in terms of changes in relative wages—to assuming that there is a single task. Table 6 reports our baseline decomposition between 1984-2003 both for the skill premium (in the left panel) and the gender wage gap (in the right panel) as well as the decompositions under the restriction that there is comparative advantage only between labor and equipment and only between labor and tasks.

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 composition and task shifters components of our decomposition go to zero. This affects the labor productivity component, since changes in labor productivity are a residual to match changes in relative wages. For instance, since the sum of the labor composition and task shifters components pushes the skill premium up, this implies that—holding fixed the importance of the capital productivity component—the strength of the labor productivity component must rise in the skill premium decomposition. Second, it implies that direct comparative advantage is the only force giving rise to the allocation of worker types to equipment types. This affects the inferred strength of worker-equipment comparative advantage, and therefore affects both the capital and labor productivity components of the decomposition. For instance, 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 strength of worker-equipment comparative advantage and increases the impact of the capital productivity component in the skill premium decomposition, thereby reducing the impact of the labor productivity component. Table 6 confirms this intuition: the strength of the capital productivity component in accounting for changes in the skill premium 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 (80%) has been driven by changes in relative equipment productivities. Similarly, because we would infer that women have a strong comparative advantage with computers in the absence of worker-task or equipment-task comparative advantage, we would therefore incorrectly conclude that changes in capital productivity played a central role in reducing the gender gap.

Abstracting from any comparative advantage at the level of equipment has similar effects. First, it implies that the capital productivity component of our decomposition goes to zero. Hence—holding fixed the importance of the labor composition and task shifters components—the absolute value of the labor productivity component must become larger in the skill premium decomposition. Second, it implies that the only force generating the allocation of worker types to tasks is 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 strength of this direct comparative advantage in the skill premium decomposition. 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 decomposition of changes in between-group inequality. It does so first by forcing changes in labor productivity to absorb the impact of the missing component(s) and second by changing the inferred strength of the remaining source of comparative advantage.

7.2 Changing comparative advantage over time

In our baseline model we imposed that the only time-varying components of productivity are multiplicatively separable between labor components, equipment components, and task components. The benefit of this assumption is that it allowed us to conduct a decomposition with four components; its cost is that it restricted comparative advantage to be fixed over time and determined solely by $T(\lambda, \kappa, \sigma)$. In practice, over time some worker types may have become relatively more productive in some tasks or using some types of equipment, and some equipment types may have become relatively more productive in some tasks. Here we generalize our baseline model to incorporate such changes over time and show that our results remain largely unchanged.

Changes in CA	Labor comp.	Task shifters	Labor prod.	Capital prod.	Labor-task	Labor-capital	Capital-task
None (baseline)	-0.084	0.119	0.099	0.067	-	-	-
Labor-capital (case 1)	-0.081	0.119	-	-	-	0.166	-
Labor-task (case 2)	-0.072	-	-	0.073	0.212	-	-
Capital-task (case 3)	-0.078	-	0.096	-	-	-	0.197

Table 7: Decomposing changes in the log skill premium allowing comparative advantage to evolve over time: 1984-2003

Specifically, we consider separately three extensions to our baseline model:[32]

$$T_t(\lambda,\kappa,\sigma) = \begin{cases} T_{\sigma t}(\sigma)\, T_{\lambda\kappa t}(\lambda,\kappa)\, T(\lambda,\kappa,\sigma) & \text{case 1} \\ T_{\kappa t}(\kappa)\, T_{\lambda\sigma t}(\lambda,\sigma)\, T(\lambda,\kappa,\sigma) & \text{case 2} \\ T_{\lambda t}(\lambda)\, T_{\kappa\sigma t}(\kappa,\sigma)\, T(\lambda,\kappa,\sigma) & \text{case 3} \end{cases}$$

In case 1, we allow for worker-equipment comparative advantage to change over time; in case 2, we allow for worker-task comparative advantage to change over time; and in case 3, we allow for equipment-task comparative advantage to change over time. In Appendix F we show how to decompose changes in between-group inequality into labor composition, task shifter, and labor-equipment components in case 1 and how to estimate the relevant transformed parameters. Details for cases 2 and 3 are available upon request. Table 7 reports our results decomposing changes in the skill premium between 1984 and 2003 in our baseline exercise as well as in cases 1, 2, and 3.

The intuition for why our results are largely unchanged is straightforward in cases 1 and 2. In each case (including case 3) our measure of changes in labor composition as well as the algorithm for conducting the labor composition decomposition are exactly the same as in our baseline model. The only change to the labor composition component is to the estimated values of factor allocations in the base period, $\{\pi_{t_0}(\lambda,\kappa,\sigma)\}$, which are inputs in the algorithm. While relaxing our assumption on $\{T_t(\lambda,\kappa,\sigma)\}$ does improve our estimates of these allocations, the improvement is small because our baseline estimates of $\{\pi_t(\lambda,\kappa,\sigma)\}$ matched the data reasonably well. Hence, the labor composition component of our baseline decomposition is robust. Similarly, our measures of changes in task productivity in case 1 and of changes in equipment productivity in case 2 as well as the algorithm for conducting the task productivity decomposition in case 1 and the

[32] If we allow for the most general form of changes in comparative advantage, $\{T_{\lambda\kappa\sigma t}(\lambda,\kappa,\sigma)\}$, then we could only decompose changes in between-group inequality into changes in the composition of the labor force and changes in overall productivity.

equipment productivity decomposition in case 2 are exactly the same as in our baseline model. Again, the only change in calculating these components is to the estimated values of factor allocation in the base period. Since these changes are small in practice, the impact of changes in task productivity in case 1 and equipment productivity in case 2 are very similar to the impacts in our baseline decomposition. Finally, since the sum of all four or three components of our decomposition (in the baseline model or the extensions considered here, respectively) closely match the change in the data in each measure of between-group inequality, we know that the sum of the labor productivity and equipment productivity components in our baseline must closely match the labor-equipment component in case 1; similarly, the sum of the labor productivity and task productivity components in our baseline must closely match the labor-task component in case 2.

Note that the robustness of our baseline exercise to changes in comparative advantage over time is not a general result that must apply to other time periods or countries. It requires that the estimated values of factor allocations in the base period are not substantially and systematically affected by allowing for changes in interactions over time.

7.3 Alternative parameter values

In this section we vary θ and ρ—recall that our decomposition results are independent 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 composition, 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 $\left\{\tau_{\lambda t}\left(\lambda\right)/\tau_{\lambda t}\left(\lambda_1\right)\right\}$ is estimated to match relative wages.

Table 8 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 8 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 group inequality—is robust to alternative values of θ within the range of $2 \leq \theta \leq 4$. At

value of θ	Labor comp.	Task shifters	Labor prod.	Capital prod.
$\theta = 2$	-0.119	0.162	0.063	0.093
$\theta = 2.9$	-0.089	0.125	0.094	0.071
$\theta = 3.1$	-0.084	0.119	0.099	0.067
$\theta = 3.3$	-0.080	0.114	0.104	0.064
$\theta = 4$	-0.068	0.098	0.117	0.055

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

$\theta = 2$ labor productivity explains only 20% of the forces increasing the skill premium between 1984 and 2003 whereas at $\theta = 4$ it explains a much larger share, but continues to explain less of the rise in the skill premium (43%) 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 composition 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 composition 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 $\left\{ \hat{\tau}_{\sigma} \left(\sigma \right) / \hat{\tau}_{\sigma} \left(\sigma \right) \right\}$, which are lower for skill-intensive occupations, are independent of ρ. Since what matters for the impact of task shifters is $\frac{\hat{\zeta}(\sigma)}{\hat{\zeta}(\sigma_1)} \left(\frac{\hat{\tau}_{\sigma}(\sigma)}{\hat{\tau}_{\sigma}(\sigma_1)} \right)^{(\rho-1)(1-\alpha)/\theta}$ for all σ, a lower value of ρ is similar to increasing task shifters in skill-intensive occupations.

Table 9 confirms this intuition. The middle row of the left panel replicates our baseline results. The left panel of Table 9 provides our decomposition results, in our baseline specification, for three values of $\rho = \{1/2, 1, 2\}$, and shows that our result—that the

	CA fixed over time				Changing capital-task CA		
value of ρ	Labor comp.	Task shifters	Labor prod.	Capital prod.	Labor comp.	Capital-task	Labor prod.
$\rho = 1/2$	-0.102	0.158	0.092	0.047	-0.094	0.222	0.089
$\rho = 1$	-0.084	0.119	0.099	0.067	-0.078	0.197	-0.120
$\rho = 2$	-0.062	0.070	0.108	0.091	-0.058	0.166	0.104

Table 9: Decomposing changes in the skill premium for alternative values of ρ: 1984-2003

combination of changes in task shifters and capital productivity explains the majority of the rise in between-education group inequality—is robust to alternative values of ρ within this range. Specifically, the the combination of these two components explains between roughly 60% (when $\rho = 2$) and 69% (when $\rho = 1/2$) of the sum of the forces pushing the skill premium upwards. 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 relatively more important.

Our extension in Section 7.2 in which CA between capital and tasks changes over time, case 3, provides an even clearer way of showing that the combination of changes in task shifters and capital productivity explains the majority of the rise in between-education group inequality for a wide range of values of ρ. In this extension, the capital and task components of the decomposition are combined into a single term. These results are shown in the right panel of Table 9. In this case, the capital-task component explains between roughly 61% (when $\rho = 2$) and 71% (when $\rho = 1/2$) of the sum of the forces pushing the skill premium upwards.

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

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

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

$$Y_{n,t}\left(C\right) = \sum_{i} d_{ni,t}\left(C\right) C_{ni,t}$$

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,

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

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

$$Y_{n,t}\left(\kappa\right) = \sum_{i} d_{ni,t}\left(\kappa\right) I_{ni,t}\left(\kappa\right)$$

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

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

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

$d_{ni,t}(\kappa) = \infty$ 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 absorption 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 $\left\{\hat{\tau}_{n,\kappa}(\kappa)/\hat{\tau}_{n,\kappa}(\kappa_1)\right\}$ 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 $\hat{s}_{nn}(\kappa)$,

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

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 parameters 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. *$\left\{\hat{w}^A_{n,t}(\lambda)/\hat{w}^A_{n,t}(\lambda_1)\right\}$ is the solution to equations (9)-(12) and (22) with $\hat{T}(\lambda,\kappa,\sigma) = \hat{L}(\lambda) = \hat{\mu}(\sigma) = \hat{q}(\kappa) = 1$ for all λ,κ,σ 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.[33]

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 Computing Machinery) and 34-35 (Transport Equipment). We obtain similar domestic absorption 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 $s_{nn,03}(\kappa_2) = 0.650$.

[33] 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 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 $\left[\frac{w_n\left(\lambda;\Phi_{t'},\Phi^*_{t'},d_{t'}\right)}{w_n\left(\lambda;\Phi_t,\Phi^*_t,d_t\right)}\right]\left[\frac{w_n\left(\lambda;\Phi_{t'},\Phi^*_{t'},d^A_{n,t'}\right)}{w_n\left(\lambda;\Phi_t,\Phi^*_t,d^A_{n,t}\right)}\right]^{-1}$. 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.

Year	HSG/HSD	SMC/HSD	CLG/HSD	GTC/HSD	Skill premium
84	-0.000	-0.000	-0.001	-0.001	-0.001
03	-0.009	-0.018	-0.024	-0.025	-0.014

Table 10: The impact on log relative wages of moving to autarky in 1984 and 2003

Table 10 reports the effect on log relative wages of moving to autarky in 1984 and 2003. The impact in 1984 is small, roughly a 0.1 percentage point decrease in the skill premium and correspondingly small changes in more disaggregated measures. The impact in 2003 is an order of magnitude larger than in 1984. Moving from 2003 to autarky generates a 1.4 percentage point reduction in the skill premium and a 2.5 percentage point reduction in the relative wage of the most educated to the least 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 10. 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 1.3 percentage points less than it did over this time period. This accounts for about 19% of the total impact of changes in capital productivity displayed in Table 4.

9 Conclusions

In this paper we provide a framework with multiple worker types, equipment types, and tasks to decompose changes in aggregated and disaggregated measures of between-group inequality into changes in (i) the composition of the workforce across labor types, (ii) the importance of different tasks, (iii) the extent of computerization, and (iv) other labor-specific productivities. The model features three forms of comparative advantage: between worker types and equipment, worker types and tasks, and equipment and tasks. We parameterize the model to match observed factor allocation and wages in the United States between 1984 and 2003 and show that the combination of changes in computerization and the relative importance of tasks explain the majority of the rise in the skill premium and the rise in inequality across more disaggregated education types as well as almost half the fall in the gender wage gap.

In spite of its high dimensionality, our framework remains empirically tractable, lending itself to a variety of extensions and applications. For instance, we extend the model to incorporate trade in consumption and capital goods and show that international trade accounts for about one fifth of the impact of computerization on the skill premium in the United States. It would be interesting to extend the model to incorporate international trade in tasks. The challenge in implementing such an extension—when mapping tasks in the model to occupations in the data, as we do here—is the lack of available data on trade in tasks.

Our framework, estimation strategy, and decomposition algorithms could be used more broadly in any country with sufficiently rich data on worker allocation to equipment types and occupations (or sectors). For instance, in ongoing work, we study the evolution of between-group inequality in Germany.

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. and David Dorn, "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, August 2013, *103* (5), 1553–97.

_ **, 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.

Katz, Lawrence F and Kevin M Murphy, "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, February 1992, *107* (1), 35–78.

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 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) $\left\{L_t(\lambda)/L_t(\lambda_1)\right\}$; (iii) $\{\psi\tau(\lambda,\kappa,\sigma)\}$ for an arbitrary constant ψ; and (iv) $\left\{\tau_{xt}(x)/\tau_{\lambda t}(x_1)\right\}$ for all $x \in \{\lambda,\kappa,\sigma\}$.

Given (i), (iii), and (iv) we construct $\{\pi_{t_0}(\lambda,\kappa,\sigma)\}$ and $\left\{w_{t_0}(\lambda)/w_{t_0}(\lambda_1)\right\}$ using equations (4), (5), and (14a) 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

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

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

Given $(i)-(iv)$, we therefore 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, $\left\{\hat{L}(\lambda)/\hat{L}(\lambda_1)\right\}$, and values of $\left\{\frac{w_{t_0}(\lambda)L_{t_0}(\lambda)}{w_{t_0}(\lambda_1)L_{t_0}(\lambda_1)}\right\}$ and $\{\pi_{t_0}(\lambda,\kappa,\sigma)\}$, changes in relative wages between t_0 and t_1 generated by changes in labor supplies 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* $\left\{\frac{w_{t_0}(\lambda)L_{t_0}(\lambda)}{w_{t_0}(\lambda_1)L_{t_0}(\lambda_1)}\right\}$ *and* $\{\pi_{t_0}(\lambda,\kappa,\sigma)\}$*, changes in relative wages between* t_0 *and* t_1 *generated by changes in labor productivities 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* $\left\{\frac{w_{t_0}(\lambda)L_{t_0}(\lambda)}{w_{t_0}(\lambda_1)L_{t_0}(\lambda_1)}\right\}$ *and* $\{\pi_{t_0}(\lambda,\kappa,\sigma)\}$*, changes in relative wages between* t_0 *and* t_1 *generated by changes in capital productivities 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* $\left\{\frac{w_{t_0}(\lambda)L_{t_0}(\lambda)}{w_{t_0}(\lambda_1)L_{t_0}(\lambda_1)}\right\}$ *and* $\{\pi_{t_0}(\lambda,\kappa,\sigma)\}$*, changes in relative wages between* t_0 *and* t_1 *generated by changes in task shifters can be calculated using*

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

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

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

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

Proof. Equation (10) and the definition of $\tau_{\sigma t}\left(\sigma\right)$ imply that between time t_0 and t_1, $\hat{T}_{\sigma}\left(\sigma\right)$ and $\hat{\mu}\left(\sigma\right)$ must satisfy the following condition:

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

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 λ,κ,σ in equations (9)-(12) and imposing condition (24), we obtain the system of equations in Lemma 4. □

B Data

Sample selection. We restrict our sample by dropping workers who are younger than 17 years old, do not report positive (paid) hours worked, or are self-employed. To use the same sample in our baseline quantification, in which we map tasks in the model to occupations in the data, and in Section E, where we map tasks in the model to sectors in the data, we drop any worker whose occupation or sector cannot be matched to our occupational or sectoral concordances.

Income. We combine the years $t, t+1, t+2$ March CPS to obtain hours worked and income to match to the year t October CPS. Following Katz and Murphy (1992), Autor et al. (2008), and many others, we multiply top-coded earnings values by 1.5. Our measure of income is wage and salary income.[34] The CPS changed how it reported wage and salary income during our sample. Until the March 1987 survey, the March CPS reported total wage and salary income in a single variable. However, starting in March 1988, the CPS reported separately primary and secondary labor earnings. In this case, we first dealt with top coding of these two variables, and then summed them together. The primary labor income variable is top-coded at 99,999 from 1988 through 1995; at 150,000 from 1996 through 2002; and at 200,000 from 2003 through 2005. The secondary labor income variable is top-coded at 99,000 from 1988 through 1995; at 25,000 from 1996 through 2002; and at 35,000 from 2003 through 2005. Finally, the total salary and wage income variable was top-coded at 75,000 in 1984 and at 99,999 in 1985 and 1986.

Occupational concordance. Our concordance of occupations across time in the CPS is based on the concordance developed in Autor and Dorn (2013). Because we further aggregate their concordance (detailed below), we are able to include in our concordance a number of occupations in the 2003 October CPS that must be dropped at a more disaggregated level.

[34] Our results are robust to considering the combination of wage and salary income, business income, and farm income.

Table 11: Occupations, their characteristics, and task shifters (1984-2003)

Occupations	Task shifter	Data	Create	Guide	Repeat	Pace	Dext.	Social
		Occupation characteristics						
Executive, administrative, managerial	1.00	5.45	5.41	6.13	4.95	1.77	1.80	7.42
Management related	0.94	5.80	4.79	4.70	5.73	1.59	1.86	6.63
Professional specialty	1.26	5.33	5.81	5.00	4.74	1.80	2.38	7.56
Technicians and related support	0.99	5.34	5.11	4.20	5.96	2.38	3.15	6.12
Financial sales and related	1.15	4.78	4.88	5.48	4.95	1.69	2.50	7.27
Retail sales	0.86	3.80	4.28	3.66	5.04	2.12	2.68	6.95
Administrative support	1.08	4.22	4.21	3.70	6.40	2.11	2.49	6.54
Housekeeping, cleaning, laundry	0.64	2.38	2.31	3.09	4.32	3.07	3.09	5.00
Protective service	0.79	4.63	4.51	4.94	6.09	1.98	3.46	7.14
Food preparation and service	1.10	3.22	3.78	4.07	4.62	2.72	3.75	6.60
Health service	1.17	3.54	4.21	3.55	5.10	2.03	3.23	7.19
Building, grounds cleaning, maintenance	0.76	2.80	3.89	3.55	4.04	2.79	3.85	5.78
Personal appearance, misc. personal care and service, recreation and hospitality	0.90	3.60	5.58	4.08	4.98	1.79	3.86	7.51
Child care	0.70	2.89	5.52	4.12	3.58	1.37	2.85	7.76
Agricultural and mining	0.53	4.34	4.25	4.23	4.41	3.60	4.20	5.56
Mechanics and repairers	0.62	4.49	4.76	4.25	4.60	2.66	4.52	5.71
Construction trades	0.56	4.11	4.70	4.84	4.40	2.84	4.12	5.74
Precision production	0.77	4.32	4.89	5.14	5.18	4.29	3.68	5.95
Machine operators, assemblers, inspectors	0.50	4.19	4.20	3.85	5.05	4.49	4.08	5.04
Transportation and material moving	0.74	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.

We aggregate the occupational variable from Autor and Dorn (2013) (denoted "ad") as follows: (1) Executive, administrative, and managerial: $3 \leq ad \leq 22$; (2) Management related: $23 \leq ad \leq 37$; (3) Professional specialty: $43 \leq ad \leq 200$; (4) Technicians and related support: $203 \leq ad \leq 235$; (5) Financial sales and related: $243 \leq ad \leq 258$; (6) Retail sales: $274 \leq ad \leq 283$; (7) Administrative support: $303 \leq ad \leq 389$; (8) Housekeeping, cleaning, laundry: $405 \leq ad \leq 408$; (9) Protective service: $415 \leq ad \leq 427$; (10) Food preparation and service: $433 \leq ad \leq 444$; (11) Health service: $445 \leq ad \leq 447$; (12) Building, grounds cleaning, and maintenance: $448 \leq ad \leq 455$; (13) Personal appearance, misc. personal care and service, recreation and hospitality: $457 \leq ad \leq 467$ and $469 \leq ad \leq 472$; (14) Child care: $ad = 468$; (15) Agriculture and mining: $473 \leq ad \leq 498$ and $614 \leq ad \leq 617$; (16) Mechanics and repairers: $503 \leq ad \leq 549$; (17) Construction: $558 \leq ad \leq 599$; (18) Precision production: $628 \leq ad \leq 699$; (19) Machine operators, assemblers, and inspectors: $703 \leq ad \leq 799$; (20) Transportation and material moving: $803 \leq ad \leq 889$. See Table 11.

C 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 matrix $\{\tau(\lambda,\kappa,\sigma)\}$ (up to an arbitrary scale parameter ψ); and, (b) the relative productivities $\{\tau_{xt}(x)/\tau_{xt}(x_0)\}$ for $x \in \lambda,\kappa,\sigma$ and for every t.

In order to estimate the different components of the $\{\tau_t(\lambda,\kappa,\sigma)\}$ (see equation (14a)), we will exclusively use data on $\{w_t(\lambda)\}$, and $\{\pi_t(\lambda,\kappa,\sigma)\}$. 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 (λ,κ,σ) and t, which denotes the population average of a dummy variable taking value 1 whenever a worker of type λ uses κ in σ.[35]

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 $\iota_{2t}(\lambda)$:

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

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

Given the Law of Large Numbers, for every t and (λ,κ,σ), and for any real number $\xi > 0$, it holds 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 5. *Given equations (14a), (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 $\xi > 0$,*

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

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

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

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

From equation (25), 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), 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), we can conclude that

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

or, equivalently,

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

□

Lemma 6. *Given Lemma 5, 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,

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

□

Lemma 7. *Given Lemma 5, 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,

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

□

Lemma 8. *Given Lemmas 1, 2, and 3, equations (17), (20), (25b) and (26b), 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:

$$\ln(\delta_{\lambda t}(\lambda))=\theta\ln\frac{w_t^*(\lambda)}{w_t^*(\lambda_1)}-\ln\frac{\sum_{\kappa,\sigma}\delta_{\kappa t}(\kappa)\delta_{\sigma t}(\sigma)\delta(\lambda,\kappa,\sigma)}{\sum_{\kappa',\sigma'}\delta(\kappa')\delta(\sigma')\delta(\lambda_1,\kappa',\sigma')}. \tag{34}$$

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

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

From Lemmas 5, 6, and 7 and equation (26b), we conclude that

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

□

Lemma 9 provides an alternative estimate of $\delta_{\lambda t}(\lambda)$, which we denote by $\delta_{\lambda t}^{Alt}(\lambda)$, to the version used in the baseline exercises.

Lemma 9. *Given Lemma 5, equations (14a), (17), (25a), and (26a), we can define an estimator $\delta^{Alt}(\lambda)$ such that, for every t, every λ, and any real number $\xi>0$,*

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

Proof. For every λ, we define our estimator $\delta^{Alt}(\lambda)$ as:

$$\ln(\delta_{\lambda t}^{Alt}(\lambda))=\frac{1}{n_\Sigma n_K}\sum_{\sigma,\kappa}\ln\frac{\pi_t^*(\lambda,\kappa,\sigma)}{\delta(\lambda,\kappa,\sigma)}\frac{\delta(\lambda_1,\kappa,\sigma)}{\pi_t^*(\lambda_1,\kappa,\sigma)}. \tag{36}$$

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

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

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

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

From Lemma 5 and equation (26a), we conclude that

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

and, using equation (14a) 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

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

or, equivalently,

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

□

Corollary 2. *Given Lemmas 5, 6, 7, 8, and 9 the estimators $\delta(\lambda,\kappa,\sigma)$, $\delta_{\lambda t}(\lambda)$, $\delta_{\lambda t}^{Alt}(\lambda)$, $\delta_{\kappa t}(\kappa)$, and $\delta_{\sigma t}(\sigma)$ satisfy, for every t, every (λ,κ,σ), and any real number $\xi > 0$,*

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

Lemma 10. *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 3. *Given equations (30) and (32), and Lemma 10, 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.*

D 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 parentheses): (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.

E Mapping tasks to sectors rather than occupations

While our baseline estimation procedure uses data on the share of workers of each type and in each occupation who report that they use computers, it does not use data on the allocation of disaggregated capital equipment to occupations. Although imperfect, such data does exist at the level of the sector (although not at the level of the occupation): the BEA reports detailed estimates for private nonresidential fixed assets by detailed industry and by detailed asset type.[36] In this section we map tasks in the model to sectors in the data in order to compare the model's implied equipment allocation across sectors to the data. To facilitate this analysis, we use BEA industry codes and aggregate up to to forty seven sectors.

Whereas mapping tasks to sectors in the data provides an opportunity to check that the model's implied equipment allocation across tasks is reasonable, there is a cost associated with mapping tasks to sectors. Relative to occupations, sectors aggregate over jobs with disparate task requirements. This has strong implications for the results of our decomposition exercises that make us more comfortable mapping tasks in the model to occupations in the data, as described in depth below.

Our decomposition and parameterization approach is identical here to our baseline. Hence, we simply report three key results: the extent to which the model's implied equipment allocation across sectors matches that in the data as well as the decomposition of changes in the skill premium and the gender gap over the full sample.

Capital equipment allocation across sectors. We construct the ratio of the share of the value of

[36]The BEA provides the following note associated with this data: "The Bureau of Economic Analysis (BEA) does not publish these detailed estimates in the SURVEY OF CURRENT BUSINESS or in the standard fixed asset tables because they are less reliable than the higher level aggregates in which they are included. Compared to these aggregates, the more detailed estimates are more likely to be based on judgmental trends, on trends in the higher level aggregate, or on less reliable source data."

	Data	Labor comp.	Task shifters	Labor prod.	Capital prod.
Skill premium	0.209	-0.034	0.044	0.086	0.117
Gender gap	-0.137	0.019	-0.069	-0.052	-0.039

Table 12: Decomposing changes in the log skill premium and log gender gap when mapping tasks in the model to sectors in the data: 1984-2003

computers relative to the share of non-computer equipment allocated to sector σ for each σ both in the data and model, and calculate the correlation between these ratios for each year in our sample. Taking the median of this correlation across years, we obtain 0.58. Hence, in spite of the fact that our estimation does not make use of data on the allocation of capital across sectors or occupations, the model's predictions line up reasonably well with the limited data that is available at the sector level.

Decomposing changes in the skill premium and the gender gap. Table 12 displays our decomposition results over the full sample for the skill premium and the gender gap when mapping tasks in the model to sectors in the data. Changes in labor composition, task shifters, and labor productivity generate smaller movements in between-group inequality than in our baseline, whereas changes in capital productivity generate larger movements. The logic is straightforward, and demonstrates the benefit of mapping tasks to occupations.

In practice, heterogeneity across jobs within a sector is much greater than within an occupation. This is true even though our occupations, of which we have 20, are more aggregated than our sectors, of which we have 47. We infer weaker worker-task comparative advantage if, in the data, employment across tasks is more similar for different worker types, as it is when we map tasks to sectors. This implies that changes in task shifters and labor composition have smaller effects on between-group inequality when we map tasks to sectors. On the other hand, mapping tasks to sectors has the opposite effect on the importance of changes in capital productivity. As an example, the large gap between men and women in their aggregate computer usage (see Table 1) remains large when conditioning on sector of employment (within sectors, women tend to work in jobs in which all workers are relatively more likely to use computers), whereas it becomes significantly smaller when conditioning on occupation of employment.

In spite of these differences, our central results are largely robust: the combination of changes task shifters and capital productivity play a central role in explaining the evolution of between-education and between-gender group inequality, in spite of the fact that these changes are not estimated off of changes in wages.

F Overview of changing comparative advantage

F.1 Decomposition exercises

Here we show how to conduct our decomposition between $t = t_0, t_1$ in case 1, in which

$$T_t(\lambda,\kappa,\sigma) = T_{\sigma t}(\sigma)\, T_{\lambda\kappa t}(\lambda,\kappa)\, T(\lambda,\kappa,\sigma)\,. \tag{38}$$

The algorithms in cases 2 and 3 can be derived similarly. Combining equations (3) and (38), we obtain

$$\tau_t(\lambda,\kappa,\sigma) = \tau_{t\lambda}(\lambda)\tau_{t\sigma}(\sigma)\tau_{t\lambda\kappa}(\lambda,\kappa)\tau(\lambda,\kappa,\sigma), \tag{39}$$

where

$$\begin{aligned}
\tau_{t\lambda}(\lambda) &= \left[\frac{T_{t\lambda\kappa}(\lambda,\kappa_1) r_t(\kappa_1)^{\frac{-\alpha}{1-\alpha}}}{\prod_{t'} T_{t'\lambda\kappa}(\lambda,\kappa_1)\, r_{t'}(\kappa_1)^{\frac{-\alpha}{1-\alpha}}}\right]^{\theta}, \\
\tau_{t\sigma}(\sigma) &= \left[\frac{T_{t\sigma}(\sigma) p_t(\sigma)^{\frac{1}{1-\alpha}}}{\prod_{t'=1}^{T} T_{t'\sigma}(\sigma) p_{t'}(\sigma)^{\frac{1}{1-\alpha}}}\right]^{\theta}, \\
\tau_{t\lambda\kappa}(\lambda,\kappa) &= \left[\frac{T_{t\lambda\kappa}(\lambda,\kappa) r_t(\kappa)^{\frac{-\alpha}{1-\alpha}}}{\prod_{t'=1}^{T} T_{t'\lambda\kappa}(\lambda,\kappa) r_{t'}(\kappa)^{\frac{-\alpha}{1-\alpha}}} \frac{\prod_{t'=1}^{T} T_{t'\lambda\kappa}(\lambda,\kappa_1) r_{t'}(\kappa_1)^{\frac{-\alpha}{1-\alpha}}}{T_{t\lambda\kappa}(\lambda,\kappa_1) r_t(\kappa_1)^{\frac{-\alpha}{1-\alpha}}}\right]^{\theta}, \\
\tau(\lambda,\kappa,\sigma) &= \left[T(\lambda,\kappa,\sigma)(1-\alpha)\alpha^{\frac{\alpha}{1-\alpha}} \prod_{t'=1}^{T} T_{t'\sigma}(\sigma) p_{t'}(\sigma)^{\frac{1}{1-\alpha}} T_{t'\lambda\kappa}(\lambda,\kappa) r_{t'}(\kappa)^{\frac{-\alpha}{1-\alpha}}\right]^{\theta}.
\end{aligned}$$

Note that $\tau_{\lambda\kappa t}(\lambda,\kappa_1) = 1$ for all λ and t.

Suppose we have the following parameters: (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 $\{\tau_{t\lambda\kappa}(\lambda,\kappa)\}$. Given (i), (iii), and (iv) we construct $\{\pi_{t_0}(\lambda,\kappa,\sigma)\}$ and $\{w_{t_0}(\lambda)/w_{t_0}(\lambda_1)\}$ using equations (4), (5), and (39) as

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

and

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

Given (ii) we construct $\frac{\hat{L}(\lambda)}{\hat{L}(\lambda_1)} = \frac{L_{t_1}(\lambda)}{L_{t_0}(\lambda)} \frac{L_{t_0}(\lambda_1)}{L_{t_1}(\lambda_1)}$ for each λ and given (iv) we construct $\frac{\hat{\tau}_x(x)}{\hat{\tau}_x(x_1)} = \frac{\tau_{xt_1}(x)}{\tau_{xt_1}(x_1)} \frac{\tau_{xt_0}(x_1)}{\tau_{xt_0}(x)}$ for each $x \in \{\lambda,\sigma\}$ and $\hat{\tau}_{\lambda\kappa}(\lambda,\kappa) = \frac{\tau_{\lambda\kappa t_1}(\lambda,\kappa)}{\tau_{\lambda\kappa t_0}(\lambda,\kappa)}$ for each (λ,κ). Hence, given $(i)-(iv)$, we therefore 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 decomposition exercise in case 1 using the previously constructed variables. The algorithms to conduct the labor composition and task-shifter components of the decomposition in case 1 are exactly as in Lemmas 1 and 4, respectively. To conduct the labor-equipment decomposition exercise, we use the following Lemma.

Lemma 11. *Given changes in labor-equipment productivities, captured by* $\left\{\frac{\hat{\tau}_{\lambda}(\lambda)}{\hat{\tau}_{\lambda}(\lambda_1)}\right\}$ *and* $\left\{\frac{\hat{\tau}_{\lambda\kappa}(\lambda,\kappa)}{\hat{\tau}_{\lambda\kappa}(\lambda,\kappa_1)}\right\}$*, and values of* $\left\{\frac{w_{t_0}(\lambda)L_{t_0}(\lambda)}{w_{t_0}(\lambda_1)L_{t_0}(\lambda_1)}\right\}$ *and* $\{\pi_{t_0}(\lambda,\kappa,\sigma)\}$*, changes in relative wages between* t_0 *and* t_1 *generated by changes in labor-equipment productivities can be calculated using*

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

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

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

Lemma 11 follows directly.

F.2 Estimation overview

Here we provide an overview—similar in structure to that provided in Section 5.2—of our estimation procedure in case 1. We omit proofs of consistency given the similarity between the proofs required in case 1 and in our baseline as well as a description of the estimation strategy in cases 2 and 3 given the similarity to case 1.

We assign θ, ρ, and α exactly as in our baseline. As in our baseline, our estimation involves three steps. In the first step we estimate the parameters that determine comparative advantage, $\{\psi\tau(\lambda,\kappa,\sigma)\}$. In the second step we estimate $\{\tau_{\sigma t}(\sigma)/\tau_{\sigma t}(\sigma_1)\}$ and $\{\tau_{\lambda\kappa t}(\lambda,\kappa)\}$. In the final step we estimate $\{\tau_{\lambda t}(\lambda)/\tau_{\lambda t}(\lambda_1)\}$ for each year, using the estimates from steps one and two.

Step 1: Equation (39) gives us

$$\log \tau\left(\lambda,\kappa,\sigma\right)=\frac{1}{T}\sum_{t=1}^{T}\log \tau_{t}\left(\lambda,\kappa,\sigma\right) \tag{40}$$

for all λ,κ,σ. Together with equation (17), we obtain an estimator of $\psi\tau\left(\lambda,\kappa,\sigma\right)$.

Step 2: Equations (4) and (39) give us

$$\frac{\pi_t(\lambda,\kappa,\sigma)}{\pi_t(\lambda,\kappa,\sigma_1)}=\frac{\tau_t(\lambda,\kappa,\sigma)}{\tau_t(\lambda,\kappa,\sigma_1)}=\frac{\tau_{\sigma t}(\sigma)\tau(\lambda,\kappa,\sigma)}{\tau_{\sigma t}(\sigma_1)\tau(\lambda,\kappa,\sigma_1)},$$

or, analogously,

$$\log\frac{\tau_{\sigma t}(\sigma)}{\tau_{\sigma t}(\sigma_1)}=\log\left(\frac{\pi_t(\lambda,\kappa,\sigma)}{\pi_t(\lambda,\kappa,\sigma_1)}\frac{\tau(\lambda,\kappa,\sigma_1)}{\tau(\lambda,\kappa,\sigma)}\right).$$

Aggregating across the different values of λ and κ, we obtain

$$\log\frac{\tau_{\sigma t}(\sigma)}{\tau_{\sigma t}(\sigma_1)}=\frac{1}{n_\Lambda n_K}\sum_{\lambda,\kappa}\log\left(\frac{\pi_t(\lambda,\kappa,\sigma)}{\pi_t(\lambda,\kappa,\sigma_1)}\frac{\tau(\lambda,\kappa,\sigma_1)}{\tau(\lambda,\kappa,\sigma)}\right), \tag{41}$$

Equations (4) and (39) also give us

$$\frac{\pi_t(\lambda,\kappa,\sigma)}{\pi_t(\lambda,\kappa_1,\sigma)}=\frac{\tau_t(\lambda,\kappa,\sigma)}{\tau_t(\lambda,\kappa_1,\sigma)}=\frac{\tau_{\lambda\kappa t}(\lambda,\kappa)\tau(\lambda,\kappa,\sigma)}{\tau_{\lambda\kappa t}(\lambda,\kappa_1)\tau(\lambda,\kappa_1,\sigma)},$$

or, analogously,

$$\tau_{\lambda\kappa t}(\lambda,\kappa)=\frac{\pi_t(\lambda,\kappa,\sigma)}{\pi_t(\lambda,\kappa_1,\sigma)}\frac{\tau(\lambda,\kappa_1,\sigma)}{\tau(\lambda,\kappa,\sigma)},$$

given that $\tau_{t\lambda\kappa}(\lambda,\kappa_1)=1$, for every λ. Aggregating across the different values of σ, we obtain

$$\log\tau_{t\lambda\kappa}(\lambda,\kappa)=\frac{1}{n_\Sigma}\sum_{\sigma}\log\left(\frac{\pi_t(\lambda,\kappa,\sigma)}{\pi_t(\lambda,\kappa_1,\sigma)}\frac{\tau(\lambda,\kappa_1,\sigma)}{\tau(\lambda,\kappa,\sigma)}\right). \tag{42}$$

Step 3: Equations (5) and (39) give us

$$\log\frac{\tau_{\lambda t}\left(\lambda\right)}{\tau_{\lambda t}\left(\lambda_1\right)}=\theta\log\frac{w_t\left(\lambda\right)}{w_t\left(\lambda_1\right)}-\log\frac{\sum_{\kappa,\sigma}\tau_{\lambda\kappa t}\left(\lambda,\kappa\right)\frac{\tau_{\sigma t}(\sigma)}{\tau_{\sigma t}(\sigma_1)}\tau\left(\lambda,\kappa,\sigma\right)}{\sum_{\kappa',\sigma'}\tau_{\lambda\kappa t}\left(\lambda_1,\kappa'\right)\frac{\tau_{\sigma t}(\sigma')}{\tau_{\sigma t}(\sigma_1)}\tau\left(\lambda_1,\kappa',\sigma'\right)}. \tag{43}$$