Technology, Skill, and the Wage Structure

Nancy L. Stokey
University of Chicago and National Bureau of Economic Research

Technical change, even if it is limited in scope, can have effects that ripple throughout the economy. Here a flexible and tractable framework, with heterogeneous workers and technologies and many tasks, is used to analyze the general equilibrium effects of technical change for a limited set of tasks. The equilibria feature positively assortative matching between workers and technologies. The effects of technical change on employment, output, prices, and wages up and down the skill and technology ladders are sharply characterized. The effects of low-skill immigration, minimum-wage legislation, and international trade are also described.

I. Introduction

Technical change, even if it is limited in scope, can have employment, output, price, and wage effects that ripple through the whole economy. This paper uses a flexible and tractable framework, with heterogeneous workers and technologies and many tasks/goods, to analyze in detail the general equilibrium effects of technical change for a limited set of tasks. Technology and human capital are assumed to be complements in production, so the labor market—which is competitive—produces positively assortative matching between technologies and skills: tasks/goods with better technologies are produced by workers with more human capital. But the quantitative allocation of workers to technologies is endogenous, determined by demands for the tasks that are produced. Hence, technical change for a limited set of tasks produces changes in employment, output levels, prices, and wages for tasks and workers not directly affected.

Why is a model of this type useful? Not only do wage differentials across skill or occupational categories change over time, but even the trends shift.

I thank Daron Acemoglu, Ufuk Akcigit, Treb Allen, Marios Angeletos, Enghin Atalay, Gadi Barlevy, Marco Bassetto, Jess Benhabib, Katarina Borovickova, Paco Buera, Ariel Burstein, Raquel Fernandez, Xavier Gabaix, Larry Jones, Boyan Jovanovic, Joe Kaboski, Greg Kaplan, Rasmus Lentz, Jeremy Lise, Bob Lucas, Sydney Ludvigson, Jesse Perla, Steve Redding, Jose Scheinkman, Ananth Seshadri, Jaume Ventura, and Gianluca Violante; seminar participants at University College London, the University of Wisconsin, the Federal Reserve Bank of Philadelphia, the Becker-Friedman Institute, the Stern School, the Federal Reserve Bank of Chicago, and the Centre de Recerca en Economia Internacional (CREi); and the editors and referees of this journal for useful comments.

[*Journal of Human Capital*, 2018, vol. 12, no. 2]

As documented by Goldin and Katz (2007) and described both succinctly and accurately in their title, the wage structure in the United States has seen "narrowing, widening, and polarizing" over the past century.[1] Explanations for these trends always involve shifts in the relative supply of and demand for skill, with shifts in demand arising from technological change. What has been missing from the discussion is a unified way to analyze how technology and skill are matched that is flexible enough to accommodate all of these trends as possible outcomes. The usual model, which features two skill groups and labor-augmenting technical change for each group, can explain narrowing or widening of wage differentials (with skill-biased or unskilled-biased technical change) but is inadequate to talk about polarizing.

The model here fills this gap. It has many intermediate goods/tasks, which are combined to produce a single final good. Tasks differ in terms of their technology level, so there is a one-dimensional technology ladder, and workers differ in their human capital, so there is also a one-dimensional skill ladder. All production functions display constant returns to scale, and all markets are perfectly competitive, so firms, as such, play no role. A competitive equilibrium consists of an allocation of skill types to tasks and a supporting set of prices and wage rates. Complementarity between skill and technology implies that the equilibrium features positively assortative matching (PAM), as in Becker's (1973) classic model of partnership formation.

After an improvement in one technology, affecting a limited set of tasks, labor is reallocated across all tasks, and all prices and wage rates change. In the model here, those effects can be sharply characterized analytically and easily computed numerically.

The results are intuitively appealing. First, and unsurprisingly, output increases and price falls for tasks that are directly affected by the technical change. General equilibrium effects are never strong enough to offset the direct effect of the shock. The effects on employment depend on the elasticity of substitution across tasks and on the change in relative match quality. To assess match quality, note that because the equilibrium features PAM, the set of skill levels employed at any particular task form an interval. Call this the "skill bin" for that task.

For elasticities across tasks that exceed unity, the substitution effect works toward pulling labor into the production of tasks that are directly affected. Since the change is an improvement in technology, this effect is reinforced at the upper threshold of the affected skill bin. Hence, em-

[1] Ober (1948) documents in detail the narrowing of wage differentials over the period 1907–47; Goldin and Margo (1992) describe the rapid compression in the 1940s, followed by a slow widening in the 1950s and 1960s that accelerated in the 1970s and early 1980s; Katz and Murphy (1992) show that the college premium rose slowly in the 1960s, fell during the 1970s, and rose sharply in the 1980s; and Autor, Katz, and Kearney (2006), Machin and Van Reenen (2007), and Autor and Dorn (2013) document labor market polarization after 1980.

ployment necessarily expands to a group of more highly skilled workers. Consequently, employment falls at tasks farther up the technology ladder, so outputs decline, and prices and wages rise. The effects are stronger for tasks with technologies closer to the one enjoying the technical change and damped for tasks farther up the ladder.

At the other end of the affected skill bin, the tendency to pull more labor in is offset by the fact that the labor is a less suitable match for the newly improved technology. Either of these forces can dominate, so employment at the lower threshold can expand to less skilled workers, or it can contract. If it expands, employment falls at tasks farther down the ladder, so outputs decline, and prices and wages rise. If it contracts, outputs farther down the ladder increase. In either case, the changes are damped for more distant tasks. The direction of the change at the lower threshold depends, in part, on the level of employment at the affected and neighboring tasks.

For elasticities of substitution across tasks below unity, the previous results are mirrored and reversed. The substitution force works toward pushing labor away from tasks that are directly affected, to increase output of complementary tasks.

At the lower threshold of the affected skill bin, this effect is reinforced by the fact that the less skilled labor at this threshold has become a worse match for the newly improved technology. Hence, employment at the affected tasks contracts among less skilled workers. Consequently, employment expands for tasks farther down the ladder, and outputs rise, with damped changes for more distant tasks. Prices and wages may fall for some tasks and workers closest to those affected by the technical change.

At the upper threshold of the affected skill bin, the tendency to push labor out is offset by the fact that more highly skilled labor is a better match for the newly improved technology. Hence, employment at the affected tasks can expand or contract. If it expands, employment falls at tasks farther up the ladder, so outputs decline, and prices and wages rise. If it contracts, employment and output increase for tasks farther up the ladder. In either case, the changes are damped for more distant tasks.

In the cases where the direction of the net effect is ambiguous, the range of the affected skill bin is important in determining the sign. For elasticities that are not too close to unity and narrow skill ranges, the effects are rather symmetric up and down the ladder from the tasks affected by the technical change. For elasticities that are close to unity, the signs are ambiguous, but the magnitude of the change is likely to be small.

As noted above, firms play no role in the analysis, and even the word is (mostly) avoided. Each worker chooses how to use his labor endowment, combining it with any of the available technologies. The worker's decision can be viewed as a choice about an occupation, with his task output used in production of the single final good.

In some contexts, the distinction between human capital and technology is blurred. Here, human capital is an asset that belongs to a single worker, who is the only one able to employ it in production. Technology is a nonrival input, used by all workers producing a particular task. Framed in terms of competitive firms, the technology for producing a particular task is available to all. However, as is shown below, the equilibrium can readily be reinterpreted as one with monopolistically competitive firms and the technology as intangible capital that is the property of the producer. In either case, the fact that it is a nonrival input distinguishes it from both human and physical capital.

The vast literature on vintage capital models suggests that the distinction between new technologies and new capital is also blurred. If a new technology requires new investment for its implementation, giving it one label or the other is largely a matter of taste. Here, physical capital is ignored, so implementing improved technologies requires no investment.

The rest of the paper is organized as follows. Section II discusses the related literature. Section III presents the basic model and characterizes the competitive equilibrium. The main results are contained in Section IV, where the model is used to study the effect of technical change for one set of tasks. In particular, we ask how the labor allocation, task outputs, task prices, and wage rates change, for all tasks and workers. A sufficient condition is provided for the conclusion that "a rising tide lifts all boats," that the improvement raises wages for all workers, even those paired with technologies that are unaffected. Section V shows how the model can be used to address policy questions: the effects of a minimum wage, of immigration, and of opening to international trade. It also analyzes the effects of eliminating PAM and of eliminating technological heterogeneity. Section VI concludes. Mathematical arguments and proofs are gathered in the appendix.

II. Related Literature

The model here is related to the extensive theoretical literature on skill-biased technical change and to the literature on assignment models of the labor market.

The first models of skill-biased technical change have two types of workers, high-skill and low-skill, performing distinct and imperfectly substitutable tasks. In particular, the aggregate supplies of the two types of labor are inputs into a constant-elasticity-of-substitution (CES) production function for the single final good, with separate (factor-augmenting) technology shocks for each type. Acemoglu (2002) provides an elegant treatment of this model and studies its ability to account for some of the major trends in employment, wages, and skill premia in the United States.

Acemoglu and Autor (2011), who call it the "canonical" model, provide a nice assessment of its limitations as well as its strengths. They point to four limitations in particular. First, technical change, whether it is skill-biased or unskilled-biased, necessarily increases the wages of both groups. The model cannot produce wage declines. Second, because there are only two types of labor, it cannot explain the "polarization" in the wage structure observed in recent years, as documented in Autor, Katz, and Kearney (2006) and Autor and Dorn (2013). Third, because it does not distinguish between skills and tasks, it is inadequate for studying the impact of technical change that affects only particular tasks. And finally, it cannot explain changes in the allocation of skill groups across tasks. Dealing with the last two limitations requires a model that distinguishes between skills and tasks.

Acemoglu and Autor (2011) also describe four features that they would like to see in an alternative to the canonical model. These are an explicit distinction between skills and tasks, at least three skill groups, comparative advantage at different tasks across different skill groups, and the ability to produce conventional substitution and complementarity across skill groups. The authors go on to develop a model with three skill groups (plus capital) and many tasks, with production technologies for each task that are linear in the four inputs. The factor weights in the linear technologies are assumed to have the property that higher-skill types have a comparative advantage in higher-index tasks. A limitation of this setup is that improvements in a "technology"—a labor-augmentation coefficient in the production functions—affect only a single skill-task pair.

Relative to the framework in Acemoglu and Autor (2011), the model here has many skill groups as well as many tasks. In addition, comparative advantage arises endogenously as a consequence of the production function, which has skill and task technology as inputs.

Another strand of the literature on technical change adds physical capital as a factor of production, as in Autor, Levy, and Murnane (2003) and Autor and Dorn (2013), and sometimes uses the strong decline in capital (equipment) prices observed in the data as the technology shock, as in Krusell et al. (2000). In these models, physical capital can enter as a substitute for low-skill or routine labor, reducing its wage rate, or as a complement to high-skill labor, raising its wage rate. Thus, increasing the supply of physical capital can produce a wide variety of effects on wages and employment patterns, depending on the type of capital. The model here has no physical capital. Although it could be added, the price in terms of tractability is not clear.

The extensive literature on assignment models goes back to Roy (1950), who used a multidimensional description of ability. In an important contribution, Sattinger (1975) uses a model very similar to the one here to examine the partial-equilibrium problem of a single employer choosing what types of workers to hire to perform various tasks. This literature is nicely reviewed in Sattinger (1993). Virtually all of it is partial equilibrium, while the setup here is a general-equilibrium model.

The model here is closest to the one in Costinot and Vogel (2010). Their model, like the one here, is a general-equilibrium setup with one-dimensional heterogeneity of both workers and tasks. On the technology side, the model here is a special case of theirs. Specifically, in the model here skill and technology are inputs into a CES function with a substitution elasticity (strictly) less than unity, while in Costinot and Vogel, the production function is required only to be (strictly) log-supermodular. Thus, the function here satisfies the requirement in Costinot and Vogel, but the converse does not hold. The additional assumption brings two important advantages, however.

The first advantage is that the model here puts no restriction on the type of technology shocks that can be analyzed. The shocks studied here are limited in scope, affecting only one set of tasks. In terms of the distribution function for technologies, a "simple" shock of this type is a rightward shift over a limited range. Thus, it satisfies first-order stochastic dominance (FOSD).

Costinot and Vogel's (2010) framework allows them to look at only two types of technology shocks, skill-biased and extreme-biased. A skill-biased shift requires the relevant distribution functions to satisfy the monotone likelihood ratio property (MLRP), a stronger condition than FOSD, in general involving shifts throughout the distribution function. A technology shift in Costinot and Vogel is "extreme-biased" if there exists a threshold technology with the property that the relevant distribution functions satisfy MLRP above the threshold and the reverse condition below the threshold. Again, extreme-biased shocks can be constructed as weighted sums of simple shocks.

Second, the results in Costinot and Vogel (2010) are only about relative wage effects, while the model here delivers conclusions about wage levels as well as output and task price levels. The relationship between the shocks here and those in Costinot and Vogel are discussed in more detail in Section V.

Costinot and Vogel (2010) also look at shifts in the distribution of skill that satisfy similar restrictions, increasing either skill abundance or skill diversity, both defined using MLRP properties. The model and methods employed here could be slightly modified to study shifts in the skill distribution. Specifically, in the setup here the technology distribution is discrete, and the skill distribution is continuous. As is seen below, this assumption makes it easy to characterize analytically the effects of a small change in one technology. A model like the one here, but with discrete skill types and continuous technologies, could be used to study the effects of shifts in the supply of skills.

In summary, compared with the literature on biased technical change, the model here allows extensive heterogeneity in both skills and tasks. Compared with the assignment literature, the model here is general equilibrium. Compared with Costinot and Vogel (2010), the CES structure imposed here makes the solution to the general-equilibrium problem easy to characterize, both analytically and numerically, allowing sharper answers to a wider range of comparative statics questions.

III. The Model

In this section, the technologies are described and the competitive equilibrium is characterized.

A. *Final-Good Technology*

The final good is produced by competitive firms using intermediate goods/tasks as inputs. A task is characterized by its technology level, $x_j > 0$. There are J such levels, indexed by $j = 1, \ldots, J$, ordered so $0 < x_1 < x_2 < \ldots < x_J$. Let γ_j be the share of tasks with technology level x_j. The total number (mass) of tasks is normalized to unity.

All inputs enter symmetrically into final-good production, but demands for them differ if their prices differ. In equilibrium, price p_j is the same for all tasks with technology level x_j. Hence, demand is the same for such tasks. Let y_j denote the (common) quantity for those tasks. The final good is produced with the constant-returns-to-scale technology,

$$y_F = \left(\sum_{j=1}^{J} \gamma_j y_j^{(\rho-1)/\rho}\right)^{\rho/(\rho-1)}, \tag{1}$$

where $\rho > 0$ is the substitution elasticity. For $\rho = 1$, the technology is Cobb-Douglas.

The final-goods sector takes the prices p_j as given. As usual, input demands are

$$y_j = \left(\frac{p_j}{p_F}\right)^{-\rho} y_F, \qquad \text{all } j, \tag{2}$$

and the price of the final good is

$$p_F = \left(\sum_{j=1}^{J} \gamma_j p_j^{1-\rho}\right)^{1/(1-\rho)}. \tag{3}$$

We take the final good as numeraire throughout, indexing prices so $p_F = 1$. Input costs exhaust revenue, so there are no profits.

The analysis could be extended to include weights on tasks. Let $\{\omega_i\}_{i=1}^{I}$ be a set of values for the weights, and let σ_{ji} be the share of tasks with the (technology, weight) pair (x_j, ω_i). Then, output of the final good is

$$y_F = \left(\sum_{j=1}^{J} \sum_{i=1}^{I} \sigma_{ji} \omega_i^{1/\rho} \tilde{y}_{ji}^{(\rho-1)/\rho}\right)^{\rho/(\rho-1)},$$

where $\tilde{y}_{ji}$ is the input of a task with characteristics (x_j, ω_i). It is straightforward to show that, in this setting, prices p_j do not depend on i and that demand for each task is

$$\tilde{y}_{ji} = \omega_i y_j, \qquad \text{all } i, j,$$

where $\{p_j, y_j\}_{j=1}^{J}$, the aggregates y_F and p_F are as above, and

$$\gamma_j \equiv \sum_{i=1}^{I} \sigma_{ji}\omega_i, \qquad \text{all } j.$$

Output and employment vary with ω_i across tasks with the same technology x_j, but the wage structure in the economy depends only on the γ_j's.

B. *Differentiated-Good Technology*

Tasks are produced with heterogeneous labor, characterized by its skill level h, as the only inputs. Assume that h has a continuous distribution. Let $G(h)$, with density $g(h)$ on $H \equiv (h_{\min}, h_{\max})$ and $0 < h_{\min} < h_{\max} \leq \infty$, denote the distribution of skill across workers. The total size (mass) of the workforce is normalized to unity, and labor supply is inelastic: each worker supplies one unit.

The total output of a task depends on the size and quality of the workforce producing it, as well as its technology level x_j. In particular, if a task with technology x_j employs workers of various human capital levels, with $\ell_j(h) \geq 0$ denoting the number (density) of each type, then total output is

$$y_j = \int \ell_j(h)\phi(h, x_j)\, dh, \qquad \text{all } j,$$

where $\phi(h, x)$ is the CES function

$$\phi(h, x) \equiv \left[\omega h^{(\eta-1)/\eta} + (1-\omega)x^{(\eta-1)/\eta}\right]^{\eta/(\eta-1)}, \qquad \eta, \omega \in (0, 1). \tag{4}$$

The elasticity of substitution η between technology and human capital is assumed to be less than unity, and ω is the relative weight on human capital.

C. *Equilibrium*

An equilibrium consists of a final output level y_F, task outputs and prices $\{y_j, p_j\}_{j=1}^{J}$, a wage function $w(h)$, where $h \in H$, and an allocation of labor across technologies that satisfy the usual optimization and market-clearing conditions.

The model allows two interpretations about market structure. One is that each task is produced by competitive firms, with each firm choosing to employ skill types that minimize unit cost. In this case, competition insures that each worker is paid his marginal revenue product.[2] Alternatively, one can suppose that workers simply choose which task to produce, with each worker choosing a task—a job—that maximizes his income. In either

[2] If $\rho > 1$, profit-making firms could be introduced by assuming that each task is supplied by a unique producer. Under this assumption, the allocation of labor, output of each task, and prices would be unchanged, but wages would reduced by the factor $(\rho - 1)/\rho$, with the residual revenue going to profits.

interpretation, task prices are taken as given by the decision maker—the firm or the worker.

In principle, the allocation of labor could be quite complicated, with any technology level x_j employing workers with skill h in various disjoint intervals and with workers of a given human capital level h producing goods with different technologies x_j. This does not occur in equilibrium, and it is straightforward to see why not.

Since labor markets are competitive, the allocation of labor across technologies is efficient. And since the elasticity of the CES function ϕ is less than unity, it is log-supermodular. Hence, efficiency requires PAM: workers with higher skill h work with higher technologies x_j (Costinot 2009). Consequently, the equilibrium labor allocation is characterized by thresholds $h_{\min} = b_0 < b_1 < \ldots < b_{J-1} < b_J = h_{\max}$, where technology x_j employs workers with skill $h \in (b_{j-1}, b_j)$. We refer to the interval (b_{j-1}, b_j) as "skill bin j." An individual with human capital $h = b_j$ is indifferent between working with technologies x_j and x_{j+1}. Since the distribution function G is continuous, the set of such workers has measure zero, and they can be allocated to either bin.

Equilibrium also requires market clearing for goods and labor. Thus, the equilibrium conditions are

1. income maximization by all types of labor,

$$\begin{aligned} w(h) &\geq p_j \phi(h, x_j), \qquad \text{all } h, \\ \text{with equality if} \quad h &\in [b_{j-1}, b_j], \qquad \text{all } j; \end{aligned} \tag{5}$$

2. market clearing for tasks: $\{y_j, p_j\}_{j=1}^{J}$ satisfy equation (2), with y_F as in equation (1); and
3. labor market clearing,

$$\int_{b_{j-1}}^{b_j} \phi(h, x_j) g(h) dh = \gamma_j y_j, \qquad \text{all } j. \tag{6}$$

The first condition implies that each task is priced at unit cost, and the last says that the total productive capacity of labor with skill $h \in (b_{j-1}, b_j)$ is sufficient for production of tasks with technology x_j.

The allocation of labor within any skill bin (b_{j-1}, b_j) across tasks with technology x_j is, to some extent, indeterminate. Equilibrium determines only the output level y_j, which is the same across tasks with technology level x_j. For concreteness, we can suppose that each task is produced by skill types in the interval (b_{j-1}, b_j) in proportion to their representation in the population, but this is not required.[3]

[3] Since ϕ has constant returns to scale, the number of firms producing any task—if firm are introduced into the narrative—is indeterminate. Only the total (productivity-weighted) labor input and total output are determined in equilibrium.

To characterize the thresholds $\{b_j\}_{j=1}^{J-1}$, note that condition (5) implies

$$\frac{w'(h)}{w(h)} = \frac{\phi_h(h, x_j)}{\phi(h, x_j)}, \qquad h \in (b_{j-1}, b_j), \qquad \text{all } j. \tag{7}$$

Hence, the equilibrium wage function is piecewise continuously differentiable, with kinks at the points $\{b_j\}_{j=1}^{J-1}$.

Since workers with skill b_j are indifferent between working with technologies x_j and x_{j+1}, it follows immediately from condition (5) and equation (2) that

$$\frac{p_{j+1}}{p_j} = \frac{\phi(b_j, x_j)}{\phi(b_j, x_{j+1})}, \tag{8}$$

$$\frac{y_{j+1}}{y_j} = \left(\frac{\phi(b_j, x_{j+1})}{\phi(b_j, x_j)}\right)^{\rho}, \qquad j = 1, \ldots, J-1. \tag{9}$$

Unit cost and price are strictly decreasing in j, and output is strictly increasing: goods with better technologies have lower prices and higher sales. If $\rho > 1$ ($\rho < 1$), then total revenue is increasing in j (decreasing in j).

To characterize equilibrium, combine equations (6) and (9) to find that $\{b_j\}_{j=1}^{J-1}$ satisfy

$$\int_{b_j}^{b_{j+1}} g(h)\phi(h, x_{j+1})\, dh = \frac{\gamma_{j+1}}{\gamma_j}\left(\frac{\phi(b_j, x_{j+1})}{\phi(b_j, x_j)}\right)^{\rho} \int_{b_{j-1}}^{b_j} g(h)\phi(h, x_j)\, dh, \qquad j = 1, \ldots, J-1. \tag{10}$$

Since $\eta < 1$, the ratio $\phi(b_j, x_{j+1})/\phi(b_j, x_j)$ is strictly increasing in b_j. Therefore, since $b_0 = h_{\min}$ is given, for any conjectured b_1, the sequence $\{b_j\}_{j=2}^{J}$ defined recursively by using equation (10) is increasing in b_1. Equilibrium requires $b_J = h_{\max}$. Thus, a solution exists, and it is unique.

Define Ψ_j to be the "total productivity" of labor in the jth skill bin,

$$\Psi_j \equiv \int_{b_{j-1}}^{b_j} \phi(h, x_j) g(h)\, dh, \qquad j = 1, \ldots, J. \tag{11}$$

Then use equation (6) to write the output of each type of good as

$$y_j = \frac{1}{\gamma_j}\Psi_j, \qquad j = 1, \ldots, J, \tag{12}$$

and write equation (10) in the more symmetric form

$$\phi(b_j, x_{j+1})^{-\rho}\frac{1}{\gamma_{j+1}}\Psi_{j+1} = \phi(b_j, x_j)^{-\rho}\frac{1}{\gamma_j}\Psi_j, \qquad j = 1, \ldots, J-1. \tag{13}$$

D. Skill Allocation

To see more clearly how workers and technologies are matched, it is useful to look at potential-wage functions, like those in Neal and Rosen (2000, fig. 3.1). Let $w^{p}(h, x_j)$ denote the wage a worker with skill h would earn producing a task with technology x_j,

$$w^{p}(h, x_j) = p_j \phi(h, x_j), \qquad \text{all } h, \quad \text{all } j.$$

Figure 1 displays potential wages as a function of h for $J = 3$ technology levels.

For fixed x_j, the potential wage $w^{p}(h, x_j)$ is strictly increasing in h, so each curve is upward sloping. As a function of x_j, there are two effects. First, the price p_j is decreasing in x_j, so the intercept decreases with x_j. In addition, since ϕ is log-supermodular, a higher x_j implies a steeper slope for ϕ as a function of h. Thus, plotted against h, for various x_j values, the potential-wage functions cross. A worker's actual wage is the maximum of his potential wages, as in condition (5). Hence, the wage function $w(h)$ is defined by the upper envelope of the four curves, and the crossing points along the upper envelope are the thresholds b_j that divide the skill range into bins.

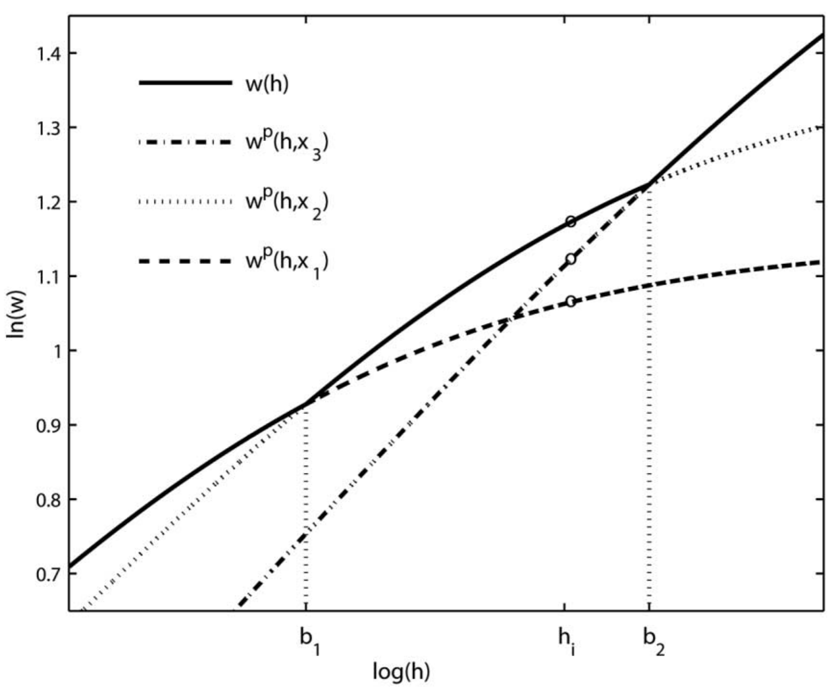

Figure 1.—Potential wages.

The three small circles in figure 1 show the choices available to a worker with skill h_i. The potential wage for that worker increases moving from x_1 to x_2, but it falls moving from x_2 to x_3, so that worker chooses x_2.

In a model with search frictions, these points would represent rungs on a job ladder for a worker with skill h_i. This worker's first job might come from an employer of any type. That firm would pay him his reservation wage, not his marginal revenue product, so his initial wage would lie below all of the displayed values. But subsequently, outside offers from other firms would raise his wage, for two reasons. If the outside firm was a better match, he would change jobs and receive a wage increase. But even if the outside firm was an equivalent (or possibly worse) match, his wage might be bid up by competition, although in this case he would not change jobs.

IV. Technical Change

This section looks at the effects of technical change that improves one technology by a small increment, with all others unchanged. Specifically, it characterizes the effect on the labor allocation, described by the thresholds $\{b_j\}_{j=1}^{J-1}$, on the output levels and prices $\{y_j, p_j\}_{j=1}^{J}$ for all tasks, and on the wage function $w(h)$.

The main forces can be previewed in figure 1. Suppose technology x_k gets the improvement. The direct effect is to increase labor productivity for workers in skill bin k, raising $w^{\mathrm{P}}(\cdot, x_k)$ and making it slightly steeper. But the higher labor productivity increases y_k, which depresses the price p_k and tends to raise all other prices, p_j, where $j \neq k$. These price changes lower $w^{\mathrm{P}}(\cdot, x_k)$ partway back toward its original level and raise all the other curves, $w^{\mathrm{P}}(\cdot, x_j)$, where $j \neq k$. The thresholds defining the employment bins shift, changing employment patterns and wages for all workers.

The rest of this section analyzes these changes in detail. Throughout, we use "hats" to denote proportionate changes induced by the perturbation, $\hat{z} \equiv z^{-1}\partial z/\partial \varepsilon$ for any variable z. All derivations and proofs are in the appendix.

A. *Final Output*

Suppose that technical change increases technology x_k by a small increment $\varepsilon > 0$, with all others unchanged. Note that the change in output of the final good y_{F} is a weighted average of the output changes for tasks,

$$\hat{y}_{\mathrm{F}} = \sum_{j=1}^{J} \nu_j \hat{y}_j, \tag{14}$$

where the weights

$$\nu_j \equiv \frac{1}{y_{\mathrm{F}}} \gamma_j p_j y_j, \qquad \text{all } j, \tag{15}$$

with $\Sigma_{j=1}^{J} \nu_j = 1$, are their cost shares in producing the final good. With the price of the final good fixed at unity, the relative price changes for tasks are

$$\hat{p}_j = \frac{1}{\rho}\left(\hat{y}_{\mathrm{F}} - \hat{y}_j\right), \qquad \text{all } j, \tag{16}$$

and the weighted average of the price changes is $\Sigma_{j=1}^{J} \nu_j \hat{p}_j = \hat{p}_{\mathrm{F}} = 0$.

Consider first the short-run effects, with labor immobile. Recall the definition of Ψ_j, for all j, in equation (11), and let $\hat{\Psi}_k$ be the direct effect of the technology improvement on total labor productivity in skill bin k. Output increases for tasks produced with technology x_k,

$$\hat{y}_k^{\mathrm{SR}} = \hat{\Psi}_k \equiv \frac{1}{\Psi_k}\frac{\partial \Psi_k}{\partial x_k} > 0, \tag{17}$$

and is unchanged for all other tasks. Hence, the change in final output is

$$\hat{y}_{\mathrm{F}}^{\mathrm{SR}} = \nu_k \hat{\Psi}_k > 0.$$

In the longer run, with labor mobile, the changes in $\{y_j\}_{j=1}^{J}$ and y_{F} must be augmented to account for the impact of changes in the skill bins, changes in the b_j's. Let $\{b_j^{(k)}(\varepsilon)\}_{j=1}^{J-1}$ denote the solution to equation (13) as a function of ε, where $b_0 = h_{\min}$ and $b_J = h_{\max}$ are fixed. Define the density-weighted changes in the thresholds

$$\beta_j^{(k)} \equiv g\left(b_j\right) b_j^{(k)\prime}(\varepsilon), \qquad j = 1, \ldots, J-1, \tag{18}$$

with $\beta_0^{(k)} = \beta_J^{(k)} = 0$. From equations (11) and (12), the long-run changes in output levels for tasks are

$$\begin{aligned} \hat{y}_k &= \frac{1}{\Psi_k}\left(\phi(b_k, x_k)\beta_k^{(k)} - \phi(b_{k-1}, x_k)\beta_{k-1}^{(k)}\right) + \hat{\Psi}_k, \\ \hat{y}_j &= \frac{1}{\Psi_j}\left(\phi\left(b_j, x_j\right)\beta_j^{(k)} - \phi\left(b_{j-1}, x_j\right)\beta_{j-1}^{(k)}\right), \qquad \text{all } j \neq k. \end{aligned} \tag{19}$$

The next proposition shows that, to a first-order approximation, the change in the labor allocation has no impact on output of the final good: the long-run increase is the same as the short-run increase.

PROPOSITION 1. In the long run, with labor mobile, the change in output of the final good is, to a first-order approximation, the same as that in the short run, $\hat{y}_{\mathrm{F}} = \hat{y}_{\mathrm{F}}^{\mathrm{SR}}$.

This result is not surprising. The potential additional effect in the long run arises only from the reallocation of labor, changes in the thresholds $\{b_j\}_{j=1}^{J-1}$ defining the skill bins. Since labor markets are competitive, the baseline allocation of labor maximizes y_{F}. Hence, to a first-order approximation, small changes in those thresholds have no effect on final output. An increase (decrease) in b_j raises (lowers) the output of tasks with technology x_j, but the effect on final output is exactly offset by the decrease (increase) in the output of tasks with technology x_{j+1}.

B. *Labor Allocation*

The changes in the labor allocation do, however, affect task-level outputs and prices, as well as wages. The rest of this section describes these changes. To determine the effect on the labor allocation, differentiate equation (13) and use equation (11) to get a system of $J - 1$ linear equations for the changes in the thresholds,

$$\underline{\beta}^{(k)} = M\underline{A}^{(k)}, \tag{20}$$

where the superscript denotes which technology has been perturbed, and for any k,

$$\begin{aligned} A_{k-1}^{(k)} &= -\rho\hat{\phi}_x(b_{k-1}, x_k) + \hat{\Psi}_k, \\ A_k^{(k)} &= \rho\hat{\phi}_x(b_k, x_k) - \hat{\Psi}_k, \\ A_j^{(k)} &= 0, \qquad \text{otherwise.} \end{aligned} \tag{21}$$

Since $\underline{A}^{(k)}$ has at most two nonzero elements—and only one if $k = 1$ or $k = J$—for fixed k the solution to equation (20) involves only $A_{k-1}^{(k)}$, $A_k^{(k)}$, and the columns $M_{\cdot k-1}$ and $M_{\cdot k}$. In particular,

$$\beta_j^{(k)} = m_{j,k-1}A_{k-1}^{(k)} + m_{j,k}A_k^{(k)}, \qquad j - 1, \ldots, J - 1, \tag{22}$$

where the first term drops out if $k = 1$ and the second drops out if $k = J$.

Here, M is the inverse of a tridiagonal matrix, so it has a recursive structure. Lemma 2 shows that it has strictly positive elements and that successive row elements above and below the diagonal have ratios that depend only on the row j.

LEMMA 2. All elements of M are positive, and the elements in each column $M_{\cdot n}$ satisfy

$$\begin{aligned} m_{j+1,n} &= q_{j+1}m_{j,n}, \qquad j \geq n, \\ m_{j-1,n} &= r_{j-1}m_{j,n}, \qquad j \leq n, \end{aligned} \tag{23}$$

where $\{q_{j+1}\}_{j=1}^{J-2}$ and $\{r_{j-1}\}_{j=2}^{J-1}$ are positive constants.

Lemma 2 can be used as follows. Fix k, and use the first line in equation (23) to compare successive rows $j + 1 > j \geq k$ in equation (20). Similarly, use the second line in equation (23) to compare successive rows $j - 1 < j \leq k - 1$, concluding that

$$\begin{aligned} \beta_{j+1}^{(k)} &= q_{j+1}\beta_j^{(k)}, \qquad j \geq k, \\ \beta_{j-1}^{(k)} &= r_{j-1}\beta_j^{(k)}, \qquad j \leq k - 1\,. \end{aligned} \tag{24}$$

Thus, all thresholds at or above the kth move in the same direction, and all those at or below the $(k-1)$th move in the same direction. It remains to determine the signs of $\beta_k^{(k)}$ and $\beta_{k-1}^{(k)}$. For this we need to characterize the two nonzero elements of $\underline{A}^{(k)}$.

Lemma 3. For any k,

1. if $\rho = 1$, then $A_{k-1}^{(k)} > 0$ and $A_k^{(k)} > 0$;
2. if $\rho > 1$, then $A_k^{(k)} > 0$ and $A_{k-1}^{(k)}$ can have either sign; and
3. if $\rho < 1$, then $A_{k-1}^{(k)} > 0$ and $A_k^{(k)}$ can have either sign.

The intuition for lemma 3 is straightforward from equation (21). The term $\hat{\phi}_x(h, x_k)$ is the proportionate change in labor productivity for a worker with skill h. Since $\eta < 1$, it is strictly increasing in h. The term $\hat{\Psi}_k$ is the average value of these changes in skill bin k. If $\rho \geq 1$, then for a worker with skill b_k at the upper threshold of the bin, $\hat{\Psi}_k < \hat{\phi}_x(b_k, x_k) \leq \rho\hat{\phi}_x(b_k, x_k)$ so $A_k^{(k)} > 0$. If $\rho < 1$, then the sign is ambiguous. Similarly, if $\rho \leq 1$, then for a worker with skill b_{k-1} at the lower threshold of skill bin k, $\rho\hat{\phi}_x(b_{k-1}, x_k) \leq \hat{\phi}_x(b_{k-1}, x_k) < \hat{\Psi}_k$, so $A_{k-1}^{(k)} > 0$. If $\rho > 1$, then the sign is ambiguous.

Can anything more be said about the terms with ambiguous signs? The answer depends, to a large extent, on how the technology levels are chosen/defined. If the technology grid is fine, then the skill bins are narrow, so b_{k-1} is close to b_k, and $A_{k-1}^{(k)} \approx -A_k^{(k)}$. For $\rho = 1$, both are close to zero.

If $A_{k-1}^{(k)}$ and $A_k^{(k)}$ are both positive, then it follows immediately from equation (22) and lemma 2 that all thresholds shift upward. But even if one term in equation (22) is negative, then the sign of the sum can sometimes be determined. Proposition 4 characterize the signs of $\beta_k^{(k)}$ and $\beta_{k-1}^{(k)}$ to the extent that it is possible.

Proposition 4. For any k, an increase in technology x_k implies

1. for $\rho = 1$,

$$\beta_j^{(k)} > 0, \qquad \text{all } j;$$

2. for $\rho > 1$,

$$\beta_j^{(k)} > 0, \qquad j \geq k,$$
$$\beta_j^{(k)} \lesseqqgtr 0, \qquad j < k-1, \qquad \text{as } \beta_{k-1}^{(k)} \lesseqqgtr 0;$$

and

3. for $\rho < 1$,

$$\beta_j^{(k)} > 0, \qquad j \leq k-1,$$
$$\beta_j^{(k)} \lesseqqgtr 0, \qquad j > k, \qquad \text{as } \beta_k^{(k)} \lesseqqgtr 0.$$

For $\rho = 1$, all thresholds shift upward. For $\rho > 1$, the thresholds at and above the kth shift upward, while those at and below the $(k-1)$th can shift either way. For $\rho < 1$, the thresholds at and below the $(k-1)$th shift upward, while those at and above the kth can shift either way.

C. Task Outputs

From equations (6) and (11), the change in output for a task of type $j \neq k$ depends on the sum of the productivity-weighted employment changes at the two thresholds,

$$\hat{y}_j^{(k)} = \frac{1}{\Psi_j}\left(\phi(b_j, x_j)\beta_j^{(k)} - \phi(b_{j-1}, x_j)\beta_{j-1}^{(k)}\right), \qquad j \neq k, \tag{25}$$

where $\beta_0^{(k)} = \beta_J^{(k)} = 0$. For goods of type k, the direct effect of the productivity change must also be added, so

$$\hat{y}_k^{(k)} = \hat{\Psi}_k + \frac{1}{\Psi_k}\left(\phi(b_k, x_k)\beta_k^{(k)} - \phi(b_{k-1}, x_k)\beta_{k-1}^{(k)}\right). \tag{26}$$

Proposition 5 characterizes the changes in output.

Proposition 5. For any k,

$$\begin{aligned} &\hat{y}_k^{(k)} > 0, \\ &\hat{y}_j^{(k)} \gtreqless 0, \qquad j > k, \qquad \text{as } \beta_k^{(k)} \lesseqgtr 0, \\ &\hat{y}_j^{(k)} \gtreqless 0, \qquad j < k, \qquad \text{as } \beta_{k-1}^{(k)} \gtreqless 0. \end{aligned}$$

Output rises for tasks of type k. The output change is in the same direction for all tasks of type $j > k$, rising if $\beta_k^{(k)} < 0$, so more labor is devoted to these tasks, and falling if $\beta_k^{(k)} > 0$. Similarly, the output change is in the same direction for all tasks of type $j < k$, falling if $\beta_{k-1}^{(k)} < 0$ and rising if $\beta_{k-1}^{(k)} > 0$. Thus, for $\rho \geq 1$, output falls for tasks of types $j > k$, and for $\rho \leq 1$, output rises for tasks of types $j < k$.

Proposition 6 shows that the size of the output changes above and below k are damped—whatever their sign—for more distant technology types.

Proposition 6. For any k,

$$\begin{aligned} &\left|\hat{y}_1^{(k)}\right| < \left|\hat{y}_2^{(k)}\right| < \ldots < \left|\hat{y}_{k-1}^{(k)}\right|, \\ &\left|\hat{y}_{k+1}^{(k)}\right| > \left|\hat{y}_{k+2}^{(k)}\right| > \ldots > \left|y_J^{(k)}\right|. \end{aligned}$$

D. Prices and Wages

Next, consider prices and wages. The price of a task rises or falls as its output change is less than or greater than the output change for the final good. In particular, from equation (16) and proposition 1,

$$\hat{p}_j^{(k)} = \frac{1}{\rho}\left(\nu_k \hat{\Psi}_k - \hat{y}_j^{(k)}\right), \qquad \text{all } j. \tag{27}$$

Proposition 7 describes price changes. For tasks of type k, price falls. For types $j \neq k$, price rises if output falls, and the size of the increase is damped for types more distant from k. The sign of the price change is ambiguous if output rises, but the price changes are nevertheless ordered, even if there is a sign change somewhere along the chain. Price decreases, if they occur, are clustered among types near k.

Proposition 7. For any k, an increase in technology x_k implies

$$\hat{p}_k^{(k)} < 0.$$

For $j < k$,

$$0 < \hat{p}_1^{(k)} < \hat{p}_2^{(k)} < \ldots < \hat{p}_{k-1}^{(k)}, \qquad \text{if } \beta_{k-1}^{(k)} < 0,$$
$$\hat{p}_{k-1}^{(k)} < \ldots < \hat{p}_2^{(k)} < \hat{p}_1^{(k)}, \qquad \text{if } \beta_{k-1}^{(k)} > 0,$$

and some or all of the latter price changes can be negative. For $j > k$,

$$0 < \hat{p}_J^{(k)} < \hat{p}_{J-1}^{(k)} < \ldots < \hat{p}_{k+1}^{(k)}, \qquad \text{if } \beta_k^{(k)} > 0,$$
$$\hat{p}_{k+1}^{(k)} < \ldots < \hat{p}_{J-1}^{(k)} < \hat{p}_J^{(k)}, \qquad \text{if } \beta_k^{(k)} < 0,$$

and some or all of the latter price changes can be negative.

Next consider wage changes. It follows immediately from condition (5) that

$$\hat{w}(h) = \hat{p}_k^{(k)} + \hat{\phi}_x(h, x_k), \qquad h \in (b_{k-1}, b_k),$$
$$\hat{w}(h) = \hat{p}_j^{(k)}, \qquad h \in \left(b_{j-1}, b_j\right), \qquad j \neq k.$$

For workers in skill bins $j \neq k$, wages change only because the price of their output changes. Hence, the direction and size of the wage change are the same as those of the price change and are equal for all workers in a skill bin. Workers in skill bin k also experience a direct productivity effect, which is increasing in the worker's own human capital h.

Proposition 8 describes the one case where a technology improvement necessarily raises all wages.

Proposition 8. If $\rho > 1$, then for any k, $\beta_{k-1}^{(k)} < 0$ implies $\hat{w}(h) > 0$, for all h.

If $\rho \leq 1$, then $\beta_{k-1}^{(k)} > 0$, leaving open the possibility that p_{k-1} falls, so wages fall for skill bin $k - 1$.

More generally, if $\beta_k^{(k)} > 0$, then workers in skill bins $j > k$ get wage increases, as do workers with human capital near the upper threshold of skill bin k. If $\beta_k^{(k)} < 0$, then wages can fall for some workers at the top of skill bin k. In this case, prices can fall for some or all tasks of type $j > k$, so that wages fall for workers in these skill bins. The wage declines are clustered near skill bin k and are damped for more distant skill bins. Indeed, wages can rise for workers sufficiently far up the skill ladder.

If $\beta_{k-1}^{(k)} < 0$, then workers in skill bins $j < k$ get wage increases, as do workers with human capital near the lower threshold of skill bin k. If $\beta_{k-1}^{(k)} > 0$, then wages can fall for some workers at the bottom of skill bin k. In this case, prices fall for some or all tasks of type $j < k$, so that wages fall for workers in these skill bins as well. The wage declines are clustered near skill bin k and are damped for more distant skill bins. Indeed, wages can rise for workers sufficiently far down the skill ladder. The appendix provides an example where wages decline for some workers.

V. Examples and Applications

In this section, we first look at several examples that illustrate the relationship between the results above and those in Costinot and Vogel (2010). The model is then used to study several substantive questions—the effects of unskilled immigration or minimum-wage legislation, of opening up to international trade, of PAM, and of technology heterogeneity.

A. Relationship to Costinot and Vogel (2010)

In the model here, the technology space is discrete and the focus is on changes in a single technology. We call such shifts “simple.”

In Costinot and Vogel (2010), the technology space is continuous, and a shift is described as a change in the density function weighting various technologies. Specifically, the technology values lie in an interval $X = [a, b]$, and their weights in the production function for the final good are represented by a continuous and strictly positive density $\gamma^{\mathrm{o}}(\cdot)$ on X. A technology shift changes the density, from γ^{o} to γ^{n}.

Costinot and Vogel study two types of shifts. A technology shift is “skill-biased” if the densities satisfy MLRP. By definition, this property holds if and only if

$$\frac{\gamma^{\mathrm{n}}(x)}{\gamma^{\mathrm{o}}(x)} \leq \frac{\gamma^{\mathrm{n}}(x')}{\gamma^{\mathrm{o}}(x')}, \qquad \text{all } x < x'.$$

That is, the ratio of the new density to the old must be (weakly) increasing in x. Lemma 5 in Costinot and Vogel shows that after such a shift, every skill type is matched to a (weakly) better technology. In addition, the two wage functions satisfy MLRP: the proportionate wage increase is larger for higher-skill workers.

A technology shift in Costinot and Vogel is “extreme-biased” if there exists a threshold technology $\hat{x}$ with the property that the densities satisfy MLRP above $\hat{x}$ and the reverse property below $\hat{x}$. Lemma 6 in Costinot and Vogel shows that after such a shift, there exists a skill threshold h^*

with the property that workers with skill above h^* are matched with better technologies and those with skill below h^* with worse technologies.

To compare the results here with those lemmas, we need to approximate discrete distributions with continuous densities and vice versa.

1. A Continuous Approximation to a Simple Shift

Fix the discrete technology levels and weights $\{x_j, \gamma_j\}_{j=1}^{J}$, and consider an increment of $\varepsilon > 0$ to technology k. For the continuous approximation, let $a = x_1 - \delta$ and $b = x_J + \delta$ (or $b = x_J + \delta + \varepsilon$, if $k = J$), where $\delta > 0$ is small, and let γ^{o} be a continuous and strictly positive approximation to $\{\gamma_j\}$. The increment to x_k is captured by a shift in the density γ^{o} to γ^{n} that replaces weight near x_k with weight near $x_k + \varepsilon$.

Clearly, such a shift never satisfies MLRP. The ratio $\gamma^{\mathrm{n}}(x)/\gamma^{\mathrm{o}}(x)$ is unity except near x_k, where it shrinks almost to zero, and near $x_k + \varepsilon$, where it explodes. Hence, no shift of this type satisfies Costinot and Vogel's definition of skill-biased. The simple shifts considered in Section IV satisfy FOSD, but not MLRP.

2. Skill-Biased Shifts

As the previous example suggests, in the discrete framework a technology shift that satisfies MLRP requires combining a series of simple shifts. One that is easy to construct is the discrete approximation to a rightward translation of the density function. Fix $X = [a, b]$ and γ^{o}. A rightward translation of γ^{o} satisfies MLRP if and only if

$$\frac{D\gamma^{\mathrm{o}}(x)}{\gamma^{\mathrm{o}}(x)} > \frac{D\gamma^{\mathrm{o}}(x')}{\gamma^{\mathrm{o}}(x')}, \qquad \text{all } x < x',$$

where $D\gamma^{\mathrm{o}} \equiv d\gamma^{\mathrm{o}}/dx$. That is, $D\gamma^{\mathrm{o}}/\gamma^{\mathrm{o}}$ must be a decreasing function. Suppose this is the case, so lemma 5 in Costinot and Vogel (2010) applies.

For the discrete approximation to γ^{o}, choose J large and let $\varepsilon = (b - a)/J$ be the size of the shift. Let $\{x_j\}_{j=1}^{J}$ be a uniform grid with step size ε, $x_1 = a + \varepsilon/2$, and $x_J = b - \varepsilon/2$. Define the probabilities $\{\gamma_j\}_{j=1}^{J}$ by

$$\gamma_j = \int_{x_j+\varepsilon/2}^{x_j-\varepsilon/2} \gamma^{\mathrm{o}}(z)\,dz, \qquad j = 1, \ldots, J.$$

In addition, let $x_{J+1} = b + \varepsilon/2$ and $\gamma_{J+1} = 0$.

Consider a rightward translation of γ^{o} by ε. In the discrete approximation, this shift changes the probabilities from $\{\gamma_j\}$ to $\{\hat{\gamma}_j\}$, defined by $\hat{\gamma}_1 = 0$, and

$$\hat{\gamma}_{j+1} = \gamma_j, \qquad j = 1, \ldots, J.$$

Moreover, this shift clearly is isomorphic to the sum of J simple shifts of the type described in Section IV.

Therefore, summing the changes in equation (20), the net effect on the thresholds is

$$\beta = M\sum_{k=1}^{J} A^{(k)} = M\begin{pmatrix} A_1^{(1)} + A_1^{(2)} \\ A_2^{(2)} + A_2^{(3)} \\ \vdots \\ A_{J-1}^{(J-1)} + A_{J-1}^{(J)} \end{pmatrix},$$

where

$$A_k^{(k)} + A_k^{(k+1)} = \rho\left(\hat{\phi}_x(b_k, x_k) - \hat{\phi}_x(b_k, x_{k+1})\right) - \left(\hat{\Psi}_k - \hat{\Psi}_{k+1}\right), \qquad \text{all } k. \quad (28)$$

As shown in the appendix, for each k, both of the terms on the right in equation (28) have order ε. Hence, the terms $A_k^{(k)} + A_k^{(k+1)}$ and the vector β also have order ε. Recall that β is defined in equation (18) as a derivative, so the vector of shifts in the thresholds induced by a rightward shift of size ε in the technology distribution is $\varepsilon\beta$. Since β itself has order ε, the vector of changes in the thresholds has order ε^2.

Does this mean that there is no task upgrading? In the setup here, no change in the thresholds means that every worker, in every skill bin, works with a technology that has improved by ε. Thus, every worker experiences task upgrading. Similarly, every absolute technology level experiences skill downgrading.

3. An Extreme-Biased Shift

For $J = 2$, a shift that is extreme-biased can be constructed from two simple shifts. Specifically, an increment of $-\varepsilon_1 < 0$ to x_1, together with an increment of $\varepsilon_2 > 0$ to x_2, satisfies the required condition. Let $\hat{b}_1$ denote the new threshold. Workers with skill below $h^* = \hat{b}_1$ experience task downgrading, and the complement experience task upgrading, in accord with lemma 6 in Costinot and Vogel (2010). Workers who remain in the same bin after the shift experience a change of size $|\varepsilon_j|$ in their technology. If $b_1' < b_1$, then workers with skill $h \in [b_1', b_1]$ experience a larger increase. If $b_1 < b_1'$, then workers with skill $h \in [b_1, b_1']$ experience a larger decrease.

B. Unskilled Immigration, Minimum-Wage Legislation

The model here can be used to study the effects of immigration by unskilled workers or minimum-wage legislation. Indeed, except for a sign change, the two policies are identical: low-skill immigration adds a segment of workers at the bottom of the skill distribution, while minimum-wage legislation subtracts a segment. The method for analyzing the effects on wages at all skill levels is exactly as in Section IV, except that the exogenous shock is a change in the supply of labor of a particular type, instead of a technology shock.

Take as a baseline the economy with the skill range $(h_{\min}, h_{\max})$. Choose $\varepsilon > 0$, and suppose that individuals with $h \in (h_{\min}, h_{\min} + \varepsilon)$ are prohibited from working, as would happen if immigration by low-skill workers is prevented or a minimum wage is imposed. Let $b_j^{(w)}(\varepsilon)$ denote the new equilibrium thresholds, as functions of ε. The lowest threshold after the change is $b_0^{(w)}(\varepsilon) = h_{\min} + \varepsilon$, while the top threshold $b_J^{(w)} = h_{\max}$ is unchanged. The (endogenous) changes in the other thresholds are determined as for a technology shock.

Formally, define $\beta_j^{(w)}$ as the slope of the function $b_j^{(w)}(\varepsilon)$, scaled by the density for skill at that point,

$$\begin{aligned} \beta_0^{(w)} &\equiv g(b_0), \\ \beta_j^{(w)} &\equiv g\big(b_j\big) b_j^{(w)\prime}, \qquad j = 1, \dots, J-1, \\ \beta_J^{(w)} &\equiv 0. \end{aligned}$$

Differentiate equation (13) and use equation (11) to get the analog of equation (20),

$$\underline{\beta}^{(w)} = M\underline{A}^{(w)},$$

where M is as before, and here the exogenous shock is the perturbation to labor supply at the bottom of the skill distribution,

$$\begin{aligned} A_1^{(w)} &\equiv \frac{1}{\Psi_1} \phi(b_0, x_1) \beta_0^{(w)} > 0, \\ A_j^{(w)} &\equiv 0, \qquad j = 2, \dots, J-1. \end{aligned}$$

Since $\underline{A}^{(w)}$ has only one nonzero element, it follows that

$$\beta_j^{(w)} = m_{j,1} A_1^{(w)}, \qquad j = 1, \dots, J-1.$$

Hence, by lemma 2, all the thresholds shift upward: $\beta_j^{(w)} > 0$, where $j = 1, \dots, J-1$.

Using equation (12) and the argument in the proof of proposition 5 (proofs for propositions 5–7 are in the appendix), the effects on task outputs are

$$\begin{aligned} \hat{y}_j^{(w)} &= -\frac{1}{\Psi_j}\left(d_j \frac{\psi_{j+1}}{\psi_j} + 1\right) \phi\big(b_{j-1}, x_j\big) \beta_{j-1}^{(w)} < 0, \qquad j = 1, \dots, J-1, \\ \hat{y}_J^{(w)} &= -\frac{1}{\Psi_J} \phi\big(b_{J-1}, x_J\big) \beta_{J-1}^{(w)} < 0, \end{aligned}$$

where $\{d_j\}_{j=1}^{J-1}$ and $\{\psi_j\}_{j=1}^{J}$ are defined in the appendix. Hence, output of every task falls. Using the argument in the proof of proposition 6,

$$\frac{\hat{y}_{j+1}^{(w)}}{\hat{y}_j^{(w)}} = -c_j \frac{\Psi_j}{\Psi_{j+1}} \frac{d_{j+1}\psi_{j+2} + \psi_{j+1}}{d_j\psi_{j+1} + \psi_j} \frac{\phi(b_j, x_{j+1})}{\phi(b_{j-1}, x_j)} < 1, \qquad j = 1, \dots, J - 1,$$

so the changes are damped farther up the task ladder,

$$|\hat{y}_1| > |\hat{y}_2| > \dots > |\hat{y}_{J-1}| > |\hat{y}_J|.$$

Under the argument in the proof of proposition 7, the proportionate price changes satisfy

$$\hat{p}_1 > \hat{p}_2 > \dots > \hat{p}_{J-1} > \hat{p}_J.$$

Since $\Sigma_1^J \nu_j \hat{p}_j = 0$, prices rise for a set of tasks at the bottom of the ladder and fall for the complementary set. That is, there exists $1 < k \le J$ such that $\hat{p}_j > 0$ for $j < k$ and $\hat{p}_j \le 0$ for $j \ge k$. Moreover, the largest price increase is at the bottom of the task ladder, for $j = 1$, with more damped changes for $1 < j < k$. For tasks at and above the kth, the price decreases are larger farther up the task ladder.

Wages changes exactly parallel the price changes. Thus, a policy that eliminates workers with skills in the range $(h_{\min}, h_{\min} + \varepsilon)$ raises wages for the remaining workers in skill bins $1 \le j < k$, where price has gone up, with larger increases farther down the skill ladder. Wages fall for workers in skill bins k and above, with larger declines farther up the skill ladder. Thus, the policy hurts the workers with $h \in (h_{\min}, h_{\min} + \varepsilon)$, who lose their jobs; benefits workers in a range just above that group, who gain from the exclusion of close competitors; and hurts workers at the upper end of the skill ladder, since relative output levels for complementary tasks fall. The workers with skill just above $h_{\min} + \varepsilon$ are the biggest winners. The minimum wage that is required to induce this shift in the baseline economy can be backed out,

$$w_{\min} = w^{(w)}(h_{\min} + \varepsilon) > w^{(\text{base})}(h_{\min} + \varepsilon) > w^{(\text{base})}(h_{\min}).$$

Figure 2 displays the results for a numerical example. The substitution elasticity between technology and skill is $\eta = 0.5$, and the two inputs have equal weight, $\omega = 0.5$. Four values are used for the elasticity of substitution across tasks, $\rho = 0.5$, 1.002, 2, and 6. The probability vector γ for technology types is a discrete approximation to a Pareto, with shape and location parameters of unity, and the skill distribution is a truncated lognormal, with mean and variance of unity, so

$$\eta = 0.5,\ \omega = 0.5,\ \rho \in \{0.5, 1.002, 2, 6\},$$
$$\lambda_F = 1,\quad x_{\min} = 1,\ \mu_h = 1, \quad \text{and} \quad \sigma_h^2 = 1.$$

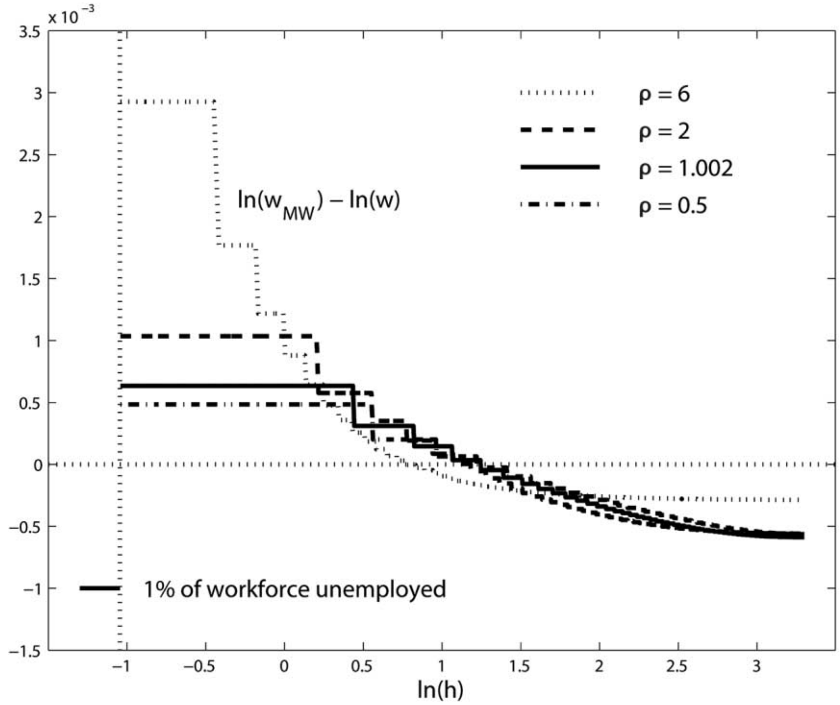

Figure 2.—Wage changes from minimum-wage legislation.

In each economy, the experiment eliminates 1 percent of the labor force at the bottom of the skill distribution. The wage function in these economies is steep at the bottom end, so in each case the wage at the first percentile is about 22 percent higher than that at the bottom, and the required minimum wage is about 22 percent higher than the lowest wage in the baseline economy. Because labor productivity is very low at the bottom of the skill distribution, the reduction in final output is small—between 0.22 percent and 0.30 percent, with lower elasticities producing larger losses.

Virtually all of the loss is borne by those who lose their jobs: the total wage bill for those who remain employed is almost unchanged. Thus, in this particular example the loss from banning immigration or from implementing a minimum wage falls almost entirely on the group that is directly excluded from employment. But in line with the theory, figure 2 shows that wages rise slightly for workers with skill in a small range above $h_m + \varepsilon$ and decline slightly for workers farther up the skill ladder.

C. *International Trade*

The effects of international trade, for two special cases, can be analyzed in a similar way. Consider a world with two countries, one large and one small, that differ only in their skill distributions, and assume that all tasks are costlessly tradeable. With costless trade, the equilibrium has a single, integrated world economy.

Suppose the large country's workers have skills in the range $[h_{\min} + \varepsilon, h_{\max})$, while the small country's workers have skills in the range $(h_{\min}, h_{\min} + \varepsilon)$. Under autarky the large country's economy is like the one in the previous section with low-skill immigrants excluded (or with a minimum wage), while the small country has a workforce that is almost homogeneous and produces the entire range of tasks. With free trade, the integrated world economy is like the one in the previous section with immigration (or without a minimum wage).

Hence, the effects on the thresholds, task outputs, prices, and wages for the large country are exactly as in the previous example. With trade, workers in the small country specialize in the lowest tasks, and all workers in the large economy experience task upgrading. World output of every task is higher in the integrated economy, but the proportionate output increases are damped farther up the task ladder. Hence, the proportionate price changes in the large country satisfy

$$\hat{p}_1 < \hat{p}_2 < \ldots < \hat{p}_{J-1} < \hat{p}_J.$$

Since the weighted price changes sum to zero, prices fall for a set of tasks $j < k$ and rise for the set $j \geq k$.

Wages in the large country follow the same pattern, falling for workers in skill bins where price has gone down and rising for those where price has increased. Trade with a skill-poor partner hurts workers in the lower part of the skill distribution, with the biggest losses at the bottom of the ladder. Trade benefits workers at the upper end of the skill ladder, since output of complementary tasks increases, with the biggest gains going to those at the top of the skill ladder.

The usual gains-from-trade argument implies that both countries enjoy increases in total output of the final good. Hence, trade benefits all the (very similar) workers in the small country. In the large country, where labor is heterogeneous, there are losers as well as winners.

D. Gain from PAM

To analyze the gains from PAM, it is useful to introduce firms and compare economies where skill is and is not observable to firms.

Suppose that each task is produced by many firms, which hire labor and sell output. Labor and task markets are competitive, so price equals unit cost for all tasks, all revenue is paid as wages, and there are no profits. Then task outputs, task prices, and the wage function are uniquely determined. The number of firms and their sizes are indeterminate but also irrelevant.

If skill is observable, then the equilibrium is exactly as before, and the economy-wide average wage is $\mathrm{E}[w(h)] = y_{\mathrm{F}}$. If skill is unobservable, then firms must hire indiscriminately and pay all workers the same wage, so in the economy with no PAM (NP), the common wage of all workers is average output, $w_{NP} = y_{FNP}$. Hence, the social gain from PAM, the increase in the average wage, is the difference in final output.

To quantify the gain, we can use a second-order approximation to the production function for final goods and, for the NP economy, approximations to the task output levels. For the NP economy, both calculations are straightforward and require no additional assumptions. For the economy with PAM, approximations are more difficult. Thus, we restrict attention to economies where technology and skill have similar distributions, so closed-form expressions are available. In addition, we require $\rho > 1$.

To approximate final output, fix a vector $(y_1, \ldots, y_J)$ of task inputs, and let $\bar{y} = \Sigma\gamma_j y_j$, $\sigma_y^2 = \Sigma\gamma_j(y_j - \bar{y})^2$, and $c_y \equiv \sigma_y/\bar{y}$ denote respectively the mean, variance, and coefficient of variation (CoV). Then final output is

$$\begin{aligned} F(y_1, \ldots, y_J) &\equiv \left(\sum_j \gamma_j y_j^{(\rho-1)/\rho}\right)^{\rho/(\rho-1)} \\ &\approx \bar{y}\left(1 - \frac{1}{\rho}\frac{1}{2}c_y^2\right), \end{aligned} \tag{29}$$

where the second line approximates around $(\bar{y}, \ldots, \bar{y})$. Thus, final output is the mean of task output, adjusted for its CoV, where the weight on the adjustment is the inverse of the substitution elasticity across tasks. Hence, the change in final output from introducing PAM has two components: the change in average task output $\bar{y}$ and the change in the CoV adjustment, the term in parentheses. As is shown below, the first is necessarily positive, but the second can have either sign. Note that the approximation is good only if the CoV of task output c_y is not too large. At a minimum, we require $c_y^2/2 < \rho$.

Fix the technology and skill distributions, and let $(\bar{x}, \sigma_x^2, c_x)$ and $(\bar{h}, \sigma_h^2, c_h)$ denote respectively the mean, variance, and CoV for each.

For the NP economy, let $q(x_j)$ denote average labor productivity at a task with technology x_j,

$$q(x_j) \equiv E_h[\phi(h, x_j)], \qquad \text{all } j.$$

Task outputs in the NP economy are

$$y_{jNP} = Z(x_1, \ldots, x_J)\, q^{\rho}(x_j), \qquad \text{all } j,$$

where

$$Z(x_1, \ldots, x_J) \equiv \left(\sum_{i=1}^{J} \gamma_i q^{\rho-1}(x_i)\right)^{-1}.$$

Hence, labor per task $Zq^{\rho-1}(x_j)$ is increasing in j. Average skill is the same across tasks, and better technologies are exploited (only) by allocating more labor to those tasks. First-order approximations to q and Z imply

$$y_{j\mathrm{NP}} \approx q(\bar{x})\left[1 + \frac{\rho q'(\bar{x})}{q(\bar{x})}\left(x_j - \bar{x}\right)\right], \qquad \text{all } j, \tag{30}$$

so task outputs have mean, variance, and CoV

$$\begin{aligned} \bar{y}_{\mathrm{NP}} &= q(\bar{x}), \\ \sigma^2_{y\mathrm{NP}} &= (pq')^2\sigma^2_x, \\ c_{y\mathrm{NP}} &= \frac{\rho\bar{x}q'}{q}\,c_x, \end{aligned} \tag{31}$$

respectively.

To analyze the equilibrium with PAM, a tractable family of economies consists of those with technology and skill distributions that jointly satisfy the following assumption.

Alignment assumption.

Let $\rho > 1$; define

$$a_H = \left[(\rho - 1)\frac{1-\omega}{\omega}\right]^{\eta/(\eta-1)},$$

let $\{x_j\}$ be a fine grid over its whole range, and let $\{\gamma_j\}$ and G together have the property that $\Sigma^j_{i=1}\gamma_i \approx G(a_H x_j)$ for all x_j.

Under the alignment assumption, skill has approximately the same distribution as technology, but scaled by a_H. Thus, $\bar{h}/\bar{x} = a_H$, and $c_h = c_x$. Let c denote their common CoV. For these economies, the competitive equilibrium with PAM has the property that both the ratio of average skill to technology, a_H, and labor per task, unity, are approximately constant across tasks.[4] Better technologies are exploited (only) by allocating labor with proportionately higher skill. Hence, task outputs are

$$y_{j\mathrm{P}} \approx \phi x_j, \qquad \text{all } j,$$

with mean, variance, and CoV

$$\begin{aligned} \bar{y}_{\mathrm{P}} &\approx \bar{x}\phi, \\ \sigma^2_{y\mathrm{P}} &\approx \phi^2\sigma^2_x, \\ c_{y\mathrm{P}} &\approx c, \end{aligned} \tag{32}$$

[4] This solution is exact if x has a continuous distribution or h has a discrete distribution.

respectively, where ϕ is evaluated at $(a_H, 1)$. Thus, the CoV for task output is the (common) CoV of the technology and skill distributions.

Under the alignment assumption, the mean and CoV of task output in the NP economy are

$$\begin{aligned} \bar{y}_{\mathrm{NP}} &\approx \bar{x}\phi\left(1 - \frac{Ac^2}{2}\right), \\ c_{y\mathrm{NP}} &\approx \frac{1 - BAc^2/2}{1 - Ac^2/2}\, c, \end{aligned} \tag{33}$$

respectively, where

$$\begin{aligned} A(\rho, \eta) &\equiv \frac{1}{\eta\rho^2}(\rho - 1) > 0, \\ B(\rho, \eta) &\equiv (\rho - 1) - \frac{1}{\eta}(\rho - 2), \end{aligned} \tag{34}$$

and B can have either sign. By definition, both q and q' are positive, which requires

$$\frac{c^2}{2} < \frac{1}{A} \quad \text{and} \quad \frac{Bc^2}{2} < \frac{1}{A}. \tag{35}$$

Recall from equation (29) that PAM affects final output by changing the mean and CoV of task output. Since ϕ is strictly concave, the effect through the mean is always positive: from equations (32) and (33),

$$\bar{y}_{\mathrm{NP}} \approx \bar{x}\phi\left(1 - \frac{Ac^2}{2}\right) < \bar{x}\phi \approx \bar{y}_{\mathrm{P}}.$$

Clearly, the size of the increase is increasing in c: a higher CoV in skill and technology increases the gain from PAM. And since A is increasing in ρ and decreasing in η, better substitutability across tasks increases the mean gain from PAM, while better substitutability between skill and technology reduces the mean gain.

The effect of PAM through the CoV of task output can have either sign. From equations (32) and (33), it reinforces or mitigates the mean effect,

$$\left(1 - \frac{1}{2}\frac{1}{\rho}\, c_{y\mathrm{P}}^2\right) \gtreqless \left(1 - \frac{1}{2}\frac{1}{\rho}\, c_{y\mathrm{NP}}^2\right),$$

or, equivalently, as $c_{y\mathrm{P}} \lesseqgtr c_{y\mathrm{NP}}$, or as

$$1 \lesseqgtr \frac{\rho\bar{x}q'}{q} \approx \frac{1 - BAc^2/2}{1 - Ac^2/2}. \tag{36}$$

For $A > 0$, the inequalities in expression (36) hold as $B \gtreqless 1$, where the approximations in equation (29) require

$$\frac{c^2}{2} < \rho / \max\left\{1, \left(\frac{1 - BAc^2/2}{1 - Ac^2/2}\right)^2\right\}. \tag{37}$$

There are several cases, depending on ρ. In all cases, good approximations to F and to q, q' require c^2 to be small. At a minimum, c^2 must satisfy expressions (35) and (37).

Case A. As $\rho \downarrow 1$, expression (34) implies $A \rightarrow 0$ and $B \rightarrow 1/\eta$, and expression (36) implies $\lim_{\rho \rightarrow 1} c_{yNP} = c_{yP}$. As the production function for final output converges to Cobb-Douglas, the CoV effect contributes nothing to the gain from PAM.

Case B. If $\rho \in (1, 2)$, then $B > 1$ and $c_{yNP} < c_{yP}$ In this case, the CoV of task output is smaller in the NP economy, mitigating the gain from PAM.

Case C. If $\rho = 2$, then $B = 1$ and $c_{yNP} = c_{yP}$. In this case, the CoV effect contributes nothing to the gain from PAM.

Case D. If $\rho > 2$, then $B < 1$ and $c_{yNP} > c_{yP}$. In this case, the CoV of task output is smaller in the economy with PAM, further increasing the gain from PAM.

Case E. As $\rho \rightarrow \infty$, expression (34) implies $A \rightarrow 0$ and $BA \rightarrow (\eta - 1)/\eta^2 < 0$, so expression (36) implies $\lim_{\rho \rightarrow \infty} c_{yNP} > c_{yP}$. As task inputs become perfectly substitutable, the CoV adjustment necessarily increases the gain from PAM.

In summary, the CoV adjustment for task output mitigates or reinforces the gain from the mean effect of PAM as $\rho < 2$ or $\rho > 2$. It is zero at $\rho = 2$, and it also vanishes as $\rho \downarrow 1$.

The elasticity of substitution η affects the magnitude of the CoV adjustment. To see this, consider the two extremes. Note from expression (36) that $\lim_{\eta \rightarrow 1} B = 1$, so as the task production function ϕ converges to Cobb-Douglas, the CoV adjustment contributes nothing to the gain from PAM. At the other extreme,

$$\lim_{\eta \rightarrow 0} B = \begin{cases} +\infty, & \rho < 2, \\ 1, & \rho = 2, \\ -\infty, & \rho > 2. \end{cases}$$

Except in the special case $\rho = 2$, as ϕ converges to Leontief, the size of the CoV adjustment diverges, with the direction of the effect depending on ρ.

E. Heterogeneous Technologies

To assess the effect of technology heterogeneity on the wage distribution, we can compare the baseline economy with one that has the same skill distribution but a single technology level. To focus on wage inequality, we choose the single technology level so that final output is the same in both economies. Then the total wage bill is also the same, and only the distribution of wages across workers changes.

With a single technology level, there are many equilibrium skill allocations, but task outputs, final output, and wages are uniquely determined. In one equilibrium, each task is produced with a pro rata share of all skill levels. For convenience, we use that one to calculate the common technology level that keeps final output unchanged.

To keep final output y_F unchanged, output of each task in the homogeneous-technology (HT) economy must be $y_{jHT} = y_{HT} = y_F$, for all j. Hence, the required technology level x_{HT} satisfies

$$y_{HT} = E_h[\phi(h, x_{HT})] .$$

Then, from equation (2) and condition (5), the wage change for a worker with human capital $h \in (b_{j-1}, b_j)$ who is in skill bin j if technologies are heterogeneous, is

$$\frac{w_{HT}(h)}{w(h)} = \frac{\phi(h, x_{HT})}{\phi(h, x_j)} \left(\frac{y_{HT}}{y_j}\right)^{-1/\rho}, \qquad h \in (b_{j-1}, b_j). \tag{38}$$

There are two forces, working in opposite directions.

In the baseline economy workers with lower skill are matched with worse technologies. Hence, a worker who in the baseline economy would be in skill bin j, with $x_j < x_{HT}$, becomes individually more productive. This effect works to raise his wage, the first term in equation (38). The reverse occurs for workers in skill bins with $x_j < x_{HT}$, so this effect works to compress the wage distribution.

But task prices also change. In the baseline economy, output y_j is increasing in x_j and price p_j is decreasing. Hence, a worker who would be in skill bin j, with output $y_j < y_{HT}$ and price $p_j > p_{HT} = 1$, suffers a cut in his product price, the second term in equation (38). Output can rise because the technology improves, because the average skill of coworkers rises, because employment rises, or any combination. This price effect, which is reversed for workers in skill bins with $y_j > y_{HT}$, works to spread the wage distribution.

The relative strength of the two forces depends on the distributions for technology and skill. For an analytical assessment it is convenient to consider economies that satisfy the alignment assumption in the previous subsection. Then in the baseline economy, all workers in skill bin j have human capital of approximately $h_j \approx a_H x_j$ and produce

$$y_j \approx \phi(h_j, x_j) = x_j \phi(a_H, 1), \qquad \text{all } j.$$

In the HT economy, they produce $\phi(h_j, x_{\mathrm{HT}})$. Hence, the wage change is

$$\begin{aligned}\Delta \ln w(h_j) &= \ln \frac{\phi(h_j, x_{\mathrm{HT}})}{\phi(h_j, x_j)} - \frac{1}{\rho}\ln\frac{y_{\mathrm{HT}}}{y_j} \\ &= \ln \frac{\phi(a_H, x_{\mathrm{HT}}/x_j)}{\phi(a_H, 1)} - \frac{1}{\rho}\ln\left(\frac{y_{\mathrm{HT}}}{x_{\mathrm{HT}}\phi(a_H,1)}\frac{x_{\mathrm{HT}}}{x_j}\right) \\ &\approx \frac{\phi_x}{\phi}\Delta_j + \frac{1}{2}\frac{\phi_x}{\phi}\left(\frac{\phi_{xx}}{\phi_x} - \frac{\phi_x}{\phi}\right)\Delta_j^2 + \chi_0 - \frac{1}{\rho}\left(\Delta_j - \frac{1}{2}\Delta_j^2\right) \\ &= \chi_0 - \frac{1}{2}\frac{\rho-1}{\rho^2}\frac{1-\eta}{\eta}\Delta_j^2,\end{aligned} \tag{39}$$

where the second line uses Euler's theorem, the third uses second-order approximations for $\phi(a_H, x_{\mathrm{HT}}/x_j)$ and $\ln(x_{\mathrm{HT}}/x_j)$, the last substitutes for ϕ_x/ϕ and ϕ_{xx}/ϕ_x, and

$$1 + \Delta_j \equiv \frac{x_{\mathrm{HT}}}{x_j}, \qquad \text{all } j,$$

$$\chi_0 \equiv -\frac{1}{\rho}\ln\left(\frac{y_{\mathrm{HT}}}{x_{\mathrm{HT}}\phi(a_H,1)}\right).$$

The coefficient on the quadratic term in equation (39) is negative, and by construction the average wage is unchanged, so $\chi_0 > 0$. Thus, workers with skill near the mean enjoy a wage gain, and those sufficiently far from the mean experience losses.

Figure 3 displays, for the same numerical examples as in figure 2, the effects of eliminating heterogeneity in technologies. In each of the four economies the common technology x_{HT} is chosen so that final output (the total wage bill) is unchanged. As shown in figure 3, the net effect in all four economies is to depress wages at both ends of the skill distribution and to raise wages for those in the middle. The loss function is approximately quadratic, as equation (39) suggests, even though here the distribution functions for skill and technology are very different from each other.

Interestingly, in every case the variance of log wages in the HT economy is slightly higher than that in the baseline economy. In these examples technology inequality reduces wage inequality, because of substantial price effects.

VI. Conclusion

The analysis here has focused on the effects of technology changes, but the framework could also be extended and used to examine other questions. As illustrated by the examples in Section V, it could be used to study

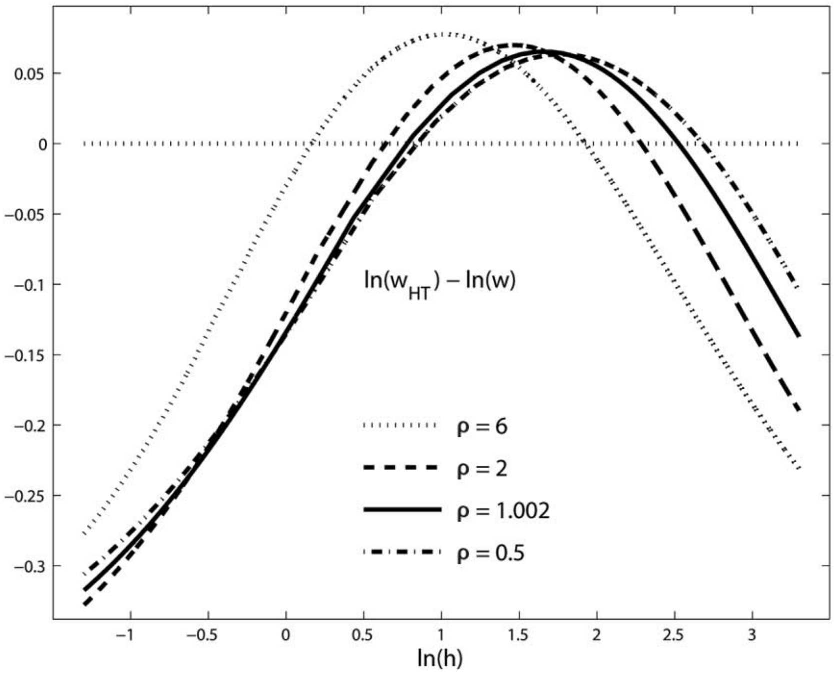

Figure 3.—Wage changes from homogeneous technologies.

the wage shifts arising from changes in immigration, minimum-wage, and trade policies.

It could also be used to revisit the role of labor market frictions in generating unemployment and producing job ladders. The model here is close to the one in Lise, Meghir, and Robin (2016), which also uses a framework with heterogeneous workers and technologies and a CES production function that combines the two inputs. Relative to that model, the one here drops search frictions but endogenizes the task prices—output prices across worker-technology pairs. Here there is a downward-sloping demand curve for each task, and its position depends on final-good production. This fact produces interactions between the wages of different workers employed at the same task and at different tasks. Closing the model in this way provides a microfoundation for the match surplus function, a function that most frictional-search models take as exogenous. As a consequence, the model here produces a nondegenerate distribution of workers across technologies/tasks, even in the absence of search frictions. Thus, it offers a richer framework for asking how important frictions are in generating wage differentials across workers.

In the framework here, individuals work in isolation to produce tasks outputs. But most goods and services, whether for consumption or investment, are not produced by single individuals. Aggregating tasks into goods requires additional information about which tasks are involved and how they are combined—a better understanding of what goes on inside firms. And since a firm may produce only one task or a wide variety of goods, these questions also require thinking about the boundaries of a firm, about the choices of which set of tasks/goods/services to sell in the marketplace, which tasks to produce in-house, and which tasks to purchase in the marketplace.

Tackling these questions is important for connecting the job/occupation decisions of individual workers with the outputs of goods/services measured in most data sources. Moreover, the patterns for recent wage changes suggest rather strongly that firms are important in determining how technical change gets translated into rising wages.[5]

Wage inequality has displayed large and long-lived shifts over the past century, as described in Goldin and Margo (1992), Goldin and Katz (2007, 2008), and Autor and Dorn (2013), and many of these shifts are surely due to changes in technology. A large increase in wage inequality leads, understandably, to calls for policies to deal with it. But to such formulate policies, we first need to better understand the underlying sources of wage inequality.

Appendix

A1. Proof of Proposition 1

Use equation (19) in equation (14) to find that

$$\hat{y}_F = \nu_k \hat{\Psi}_k + \sum_{j=1}^{J} \nu_j \frac{1}{\Psi_j} \left(\phi(b_j, x_j) \beta_j^{(k)} - \phi(b_{j-1}, x_j) \beta_{j-1}^{(k)} \right).$$

Hence, it suffices to show that

$$\begin{aligned} 0 &= \sum_{j=1}^{J-1} \left(\frac{\gamma_j p_j y_j}{\Psi_j} \phi(b_j, x_j) - \frac{\gamma_{j+1} p_{j+1} y_{j+1}}{\Psi_{j+1}} \phi(b_j, x_{j+1}) \right) \beta_j^{(k)} \\ &= \frac{\rho}{\rho - 1} \sum_{j=1}^{J-1} \left(\frac{\gamma_j y_j}{\Psi_j} - \frac{\gamma_{j+1} y_{j+1}}{\Psi_{j+1}} \right) w(b_j) \beta_j^{(k)} \\ &= \frac{\rho}{\rho - 1} \sum_{j=1}^{J-1} \left(\frac{\gamma_j \phi(b_j, x_j)^{\rho}}{\Psi_j} - \frac{\gamma_{j+1} \phi(b_j, x_{j+1})^{\rho}}{\Psi_{j+1}} \right) \phi(b_j, x_j)^{-\rho} y_j w(b_j) \beta_j^{(k)}, \end{aligned}$$

where the first line uses equation (15) and the fact that $\beta_0^{(k)} = \beta_J^{(k)} = 0$, the second uses condition (5), and the third uses equation (9). From equation (13), the term in the outer parentheses in the last line is zero, for all j. QED

[5] For example, see Song et. al. (2018).

A2. Matrix M *and Proofs of Results 2–8*

Differentiate equation (13) and use equation (11) to get

$$\begin{aligned} A_j^{(k)} &= -\frac{1}{\Psi_j}\phi(b_{j-1}, x_j)\beta_{j-1}^{(k)} \\ &\quad + \left(\frac{1}{\Psi_j}\phi(b_j, x_j) + \frac{1}{\Psi_{j+1}}\phi(b_j, x_{j+1}) + \rho\delta_j\right)\beta_j^{(k)} \\ &\quad - \frac{1}{\Psi_{j+1}}\phi(b_{j+1}, x_{j+1})\beta_{j+1}^{(k)}, \qquad j = 1, \dots, J-1. \end{aligned}$$

Write this in matrix form as

$$\underline{A}^{(k)} = T\underline{\beta}^{(k)},$$

where $\underline{A}^{(k)}$ is defined in equation (21) and T is a tridiagonal matrix of dimension $(J-1)$, with rows $(0, \dots, 0, c_j, a_j, d_{j+1}, 0, \dots, 0)$, where

$$\begin{aligned} a_j &\equiv \rho\delta_j - (d_j + c_{j+1}) > 0, \qquad j = 1, \dots, J-1, \\ \delta_j &\equiv \frac{\hat{\phi}_h(b_j, x_{j+1}) - \hat{\phi}_h(b_j, x_j)}{g(b_j)} > 0, \qquad j = 1, \dots, J-1, \\ c_j &\equiv -\frac{1}{\Psi_j}\phi(b_{j-1}, x_j) < 0, \qquad j = 2, \dots, J, \\ d_j &\equiv -\frac{1}{\Psi_j}\phi(b_j, x_j) < 0, \qquad j = 1, \dots, J-1. \end{aligned} \tag{A1}$$

The matrix in equation (20) is the inverse, $M = T^{-1}$.

To characterize M, define the constants $\{\theta_i\}_{i=0}^{J-1}$ and $\{\psi_i\}_{i=1}^{J}$ by

$$\begin{aligned} \theta_0 &\equiv 1, \\ \theta_1 &\equiv a_1, \\ \theta_i &\equiv a_i\theta_{i-1} - c_i d_i \theta_{i-2}, \qquad i = 2, \dots, J-1; \end{aligned} \tag{A2}$$

$$\begin{aligned} \psi_J &\equiv 1, \\ \psi_{J-1} &\equiv a_{J-1}, \\ \psi_i &\equiv a_i\psi_{i+1} - c_{i+1}d_{i+1}\psi_{i+2}, \qquad i = J-2, \dots, 1. \end{aligned} \tag{A3}$$

Lemma A1 shows that these constants and certain sums are positive.

LEMMA A1. The constants satisfy $\theta_i > 0$ for all i and $\psi_i > 0$ for all i, and, in addition,

$$\theta_{i-1} + c_i\theta_{i-2} > 0, \qquad i = 2, \dots, J-1, \tag{A4}$$

$$\psi_i + d_i\psi_{i+1} > 0, \qquad i = J-2, \dots, 1. \tag{A5}$$

Proof of Lemma A1. Use equation (A1) in equation (A2) to find that

$$\theta_i + c_{i+1}\theta_{i-1} = \rho\delta_i\theta_{i-1} - d_i(\theta_{i-1} + c_i\theta_{i-2}), \qquad i = 2, \ldots, J-1 .$$

Since $\rho\delta_i > 0$, $d_i < 0$, and $c_{i+1} < 0$ for all i, it follows that

$$\theta_{i-1} > 0 \quad \text{and} \quad \theta_{i-1} + c_i\theta_{i-2} > 0 \Rightarrow \theta_i + c_{i+1}\theta_{i-1} > 0 \quad \text{and} \quad \theta_i > 0 .$$

Since $\theta_1 = a_1 > 0$ and

$$\theta_1 + c_2\theta_0 = a_1 + c_2 > 0,$$

by induction expression (A4) holds. Similarly, use equation (A1) in equation (A3) to find that

$$\psi_i + d_i\psi_{i+1} = \rho\delta_i\psi_{i+1} - c_{i+1}(\psi_{i+1} + d_{i+1}\psi_{i+2}), \qquad i = J-2, \ldots, 1,$$

so

$$\psi_{i+1} > 0 \quad \text{and} \quad \psi_{i+1} + d_{i+1}\psi_{i+2} > 0 \Rightarrow \psi_i + d_i\psi_{i+1} > 0 \quad \text{and} \quad \psi_i > 0.$$

Since $\psi_{J-1} = a_{J-1} > 0$ and

$$\psi_{J-1} + d_{J-1}\psi_J = a_{J-1} + d_{J-1} > 0,$$

by induction expression (A5) holds. QED

Proof of Lemma 2. The matrix M has elements

$$\begin{aligned} m_{nn} &= \frac{1}{\theta_{J-1}}\theta_{n-1}\psi_{n+1}, \qquad n = 1, \ldots, J-1, \\ m_{j+1,n} &= -c_{j+1}\frac{\psi_{j+2}}{\psi_{j+1}} m_{j,n}, \qquad j = n, \ldots, J-2, \\ m_{j-1,n} &= -d_j\frac{\theta_{j-2}}{\theta_{j-1}} m_{j,n}, \qquad j = n, \ldots, 2. \end{aligned} \tag{A6}$$

(see Huang and McColl 1997). Since $\theta_i, \psi_i > 0$ and $d_i, c_i < 0$ for all i, clearly $m_{jn} > 0$ for all j, n. In addition, clearly the columns satisfy equation (23), where

$$\begin{aligned} q_{j+1} &\equiv -c_{j+1}\frac{\psi_{j+2}}{\psi_{j+1}}, \qquad j = 1, \ldots, J-2, \\ r_{j-1} &\equiv -d_j\frac{\theta_{j-2}}{\theta_{j-1}}, \qquad j = 2, \ldots, J. \end{aligned} \tag{A7}$$

QED

Proof of Lemma 3. From the definitions of $\hat{\phi}_x$ and $\hat{\Psi}_x$,

$$A_k^{(k)} = \rho\frac{\phi_x(b_k, x_k)}{\phi(b_k, x_k)} - \frac{\int_{b_{k-1}}^{b_k}\phi_x(h, x_k)g(h)dh}{\int_{b_{k-1}}^{b_k}\phi(h, x_k)g(h)dh},$$

and since ϕ is a CES function,

$$\phi_x(h, x) = (1-\omega)x^{-1/\eta}\phi(h, x)^{1/\eta}.$$

Hence, $A_k^{(k)} \gtreqless 0$ as

$$\int_{b_{k-1}}^{b_k} \phi(h, x_k)\big(\rho\phi(b_k, x_k)^{1/\eta-1} - \phi(h, x_k)^{1/\eta-1}\big)g(h)\,dh \gtreqless 0. \tag{A8}$$

An analogous argument (with careful attention to signs) establishes that $A_{k-1}^{(k)} \gtreqless 0$ as

$$\int_{b_{k-1}}^{b_k} \phi(h, x_k)\big(\rho\phi(b_{k-1}, x_k)^{1/\eta-1} - \phi(h, x_k)^{1/\eta-1}\big)g(h)\,dh \gtreqless 0. \tag{A9}$$

Recall that $\phi(\cdot, x_k)$ is increasing in its first argument, and $\eta < 1$. For $\rho \geq 1$, the term in the outer parentheses in expression (A8) is positive over the range of integration, so $A_k^{(k)} > 0$. For $\rho \leq 1$, the term in the outer parentheses in expression (A9) is negative, so $A_{k-1}^{(k)} > 0$. In other cases, the signs are ambiguous. QED

Proof of Proposition 4. For $\rho = 1$, the claims are immediate from equation (22) and lemmas 2 and 3. For $\rho \neq 1$, the same is true for $k = 1$ and $k = J$, since equation (22) has only one term.

For $\rho \neq 1$ and $k \neq 1, J$, use the first line of equation (23), with $j = n = k - 1$, in equation (22) to find that

$$\begin{aligned} \beta_k^{(k)} &= q_k m_{k-1,k-1} A_{k-1}^{(k)} + m_{k,k} A_k^{(k)} \\ &= \frac{\psi_{k+1}}{\theta_{J-1}}\left(-c_k\theta_{k-2}A_{k-1}^{(k)} + \theta_{k-1}A_k^{(k)}\right), \end{aligned} \tag{A10}$$

where the second line uses equations (A6) and (A7). Similarly, use the second line of equation (23), with $j = n = k$, in equation (22) to find that

$$\begin{aligned} \beta_{k-1}^{(k)} &= m_{k-1,k-1} A_{k-1}^{(k)} + r_{k-1} m_{k,k} A_k^{(k)} \\ &= \frac{\theta_{k-2}}{\theta_{J-1}}\left(-\psi_k A_{k-1}^{(k)} - d_k\psi_{k+1}A_k^{(k)}\right). \end{aligned} \tag{A11}$$

Suppose $\rho > 1$. Then $A_k^{(k)} > 0$, so the second term in equation (A10) is positive. If, in addition, $A_{k-1}^{(k)} \geq 0$, then the first term is nonnegative, so $\beta_k^{(k)} > 0$. If $A_{k-1}^{(k)} < 0$, then

$$\begin{aligned} 0 &< \int_{b_{k-1}}^{b_k} \phi(h, x_k)^{1/\eta} g(h)\,dg \\ &< \rho\phi(b_{k-1}, x_k)^{1/\eta-1}\int_{b_{k-1}}^{b_k} \phi(h, x_k)g(h)\,dh \\ &< \rho\phi(b_k, x_k)^{1/\eta-1}\int_{b_{k-1}}^{b_k} \phi(h, x_k)g(h)\,dh, \end{aligned}$$

so $|A_{k-1}^{(k)}| < A_k^{(k)}$. Hence, by lemma A1 the sum in parentheses in equation (A10) is positive. In equation (A11), the fact that $|A_{k-1}^{(k)}| < A_k^{(k)}$ does not help in applying lemma A1, so the sign is ambiguous.

Similarly, suppose $\rho < 1$. Then $A_{k-1}^{(k)} > 0$, so the first term in equation (A11) is positive. If, in addition, $A_k^{(k)} \geq 0$, then the second term is nonnegative, so $\beta_{k-1}^{(k)} > 0$. If $A_k^{(k)} < 0$, then

$$\int_{b_{k-1}}^{b_k} \phi(h, x_k)^{1/\eta} g(h)\, dg > \rho\phi(b_k, x_k)^{1/\eta - 1} \int_{b_{k-1}}^{b_k} \phi(h, x_k) g(h)\, dh$$
$$> \rho\phi(b_{k-1}, x_k)^{1/\eta-1} \int_{b_{k-1}}^{b_k} \phi(h, x_k) g(h)\, dh > 0,$$

so $|A_k^{(k)}| < A_{k-1}^{(k)}$. Hence, by lemma A1 the sum in parentheses in equation (A11) is positive. In equation (A10), the fact that $|A_k^{(k)}| < A_{k-1}^{(k)}$ does not help in applying lemma A1, so the sign is ambiguous. QED

Proof of Proposition 5. Recall from equation (A1) that

$$\phi(b_j, x_j) c_j = \phi(b_{j-1}, x_j) d_j, \qquad \text{all } j. \tag{A12}$$

For $j > k$, use the first line of equation (24) in equation (25) to find that

$$\begin{aligned} \hat{y}_j^{(k)} &= \frac{1}{\Psi_j} \left(\phi(b_j, x_j) q_j - \phi(b_{j-1}, x_j)\right) \beta_{j-1}^{(k)} \\ &= \frac{1}{\Psi_j} \left(-c_j \frac{\psi_{j+1}}{\psi_j} \phi(b_j, x_j) - \phi(b_{j-1}, x_j) \right) \beta_{j-1}^{(k)} \\ &= -\frac{1}{\Psi_j} \left(d_j \frac{\psi_{j+1}}{\psi_j} + 1 \right) \phi(b_{j-1}, x_j) \beta_{j-1}^{(k)} \\ &\gtreqless 0, \qquad \text{as } \beta_k^{(k)} \gtreqless 0, \end{aligned} \tag{A13}$$

where the second line uses the definition of q_j, the third uses equation (A12), and the last uses lemma A1 and proposition 4. Similarly, for $j < k$, use the second line in equation (24) in equation (25) and the definition of r_{j-1} to find that

$$\begin{aligned} \hat{y}_j^{(k)} &= \frac{1}{\Psi_j} \left(\phi(b_j, x_j) - r_{j-1}\phi(b_{j-1}, x_j)\right) \beta_j^{(k)} \\ &= \frac{1}{\Psi_j} \left(\phi(b_j, x_j) + d_j \frac{\theta_{j-2}}{\theta_{j-1}} \phi(b_{j-1}, x_j) \right) \beta_j^{(k)} \\ &= \frac{1}{\Psi_j} \left(1 + c_j \frac{\theta_{j-2}}{\theta_{j-1}} \right) \phi(b_j, x_j) \beta_j^{(k)} \\ &\gtreqless 0, \qquad \text{as } \beta_{k-1}^{(k)} \gtreqless 0, \qquad j < k. \end{aligned} \tag{A14}$$

For $j = k$, the first term in equation (26) is clearly positive. If $\rho \geq 1$, then the second term is also positive. If, in addition, $\beta_{k-1}^{(k)} \leq 0$, then last term is nonnegative and $\hat{y}_k^{(k)} > 0$. If $\beta_{k-1}^{(k)} > 0$, use the fact that equilibrium requires

$$\phi(b_{k-1}, x_{k-1}) p_{k-1} = \phi(b_{k-1}, x_k) p_k,$$

before and after the shock. Hence,

$$\hat{p}_{k-1} - \hat{p}_k = \left(\hat{\phi}_h(b_{k-1}, x_k) - \hat{\phi}_h(b_{k-1}, x_{k-1})\right) \frac{\beta_{k-1}^{(k)}}{g(b_{k-1})} + \hat{\phi}_x(b_{k-1}, x_k). \tag{A15}$$

For $\beta_{k-1}^{(k)} > 0$, both terms on the right are positive, so $\hat{p}_k < \hat{p}_{k-1}$. Hence, $\hat{y}_k > \hat{y}_{k-1}$, and as shown above, in this case $\hat{y}_{k-1}^{(k)} > 0$.

If $\rho < 1$, then the first and third terms in equation (26) are positive. If, in addition, $\beta_k^{(k)} \geq 0$, then the second term is nonnegative, and $\hat{y}_k^{(k)} > 0$. If $\beta_k^{(k)} < 0$, use the fact that equilibrium requires

$$\phi(b_k, x_k)p_k = \phi(b_k, x_{k+1})p_{k+1},$$

before and after the shock. Hence,

$$\hat{p}_k - \hat{p}_{k+1} = \left(\hat{\phi}_h(b_k, x_{k+1}) - \hat{\phi}_h(b_k, x_k)\right)\frac{\beta_k^{(k)}}{g(b_k)} - \hat{\phi}_x(b_k, x_{k+1}).$$

For $\beta_{k-1}^{(k)} < 0$, both terms on the right are negative, so $\hat{p}_k < \hat{p}_{k+1}$. Hence, $\hat{y}_k > \hat{y}_{k+1}$, and as shown above, in this case $\hat{y}_{k+1}^{(k)} > 0$. QED

Proof of Proposition 6. For $j > k$, use equation (A13), the fact that $\beta_j^{(k)}/\beta_{j-1}^{(k)} = q_j$, and the definition of q_j to find that

$$\begin{aligned}
\frac{\hat{y}_{j+1}^{(k)}}{\hat{y}_j^{(k)}} &= -\frac{c_{j+1}}{c_j}\frac{d_{j+1}\psi_{j+2}/\psi_{j+1} + 1}{d_j\psi_{j+1}/\psi_j + 1}\frac{\psi_{j+1}}{\psi_j}c_j \\
&= \frac{-c_{j+1}\left(d_{j+1}\psi_{j+2} + \psi_{j+1}\right)}{d_j\psi_{j+1} + a_j\psi_{j+1} - d_{j+1}c_{j+1}\psi_{j+2}} \\
&= \frac{-c_{j+1}\left(d_{j+1}\psi_{j+2} + \psi_{j+1}\right)}{\rho\delta_j\psi_{j+1} - c_{j+1}\left(\psi_{j+1} + d_{j+1}\psi_{j+2}\right)} < 1, \qquad j > k,
\end{aligned}$$

where the second line uses the definitions of ψ_j, the third uses the definition of a_j, and the inequality follows from lemma A1 and the fact that $c_{j+1} < 0$.

Similarly, for $j < k$, use equation (A14), the fact that $\beta_{j-1}^{(k)}/\beta_j^{(k)} = r_{j-1}$, and the definitions of r_{j-1}, θ_{j-1}, and a_{j-1} to find that

$$\begin{aligned}
\frac{\hat{y}_{j-1}^{(k)}}{\hat{y}_j^{(k)}} &= -\frac{d_{j-1}}{d_j}\frac{1 + c_{j-1}\theta_{j-3}/\theta_{j-2}}{1 + c_j\theta_{j-2}/\theta_{j-1}}\frac{\theta_{j-2}}{\theta_{j-1}}d_j \\
&= \frac{-d_{j-1}\left(\theta_{j-2} + c_{j-1}\theta_{j-3}\right)}{a_{j-1}\theta_{j-2} - d_{j-1}c_{j-1}\theta_{j-3} + c_j\theta_{j-2}} \\
&= \frac{-d_{j-1}\left(\theta_{j-2} + c_{j-1}\theta_{j-3}\right)}{\rho\delta_{j-1}\theta_{j-2} - d_{j-1}\left(\theta_{j-2} + c_{j-1}\theta_{j-3}\right)} < 1, \qquad j < k.
\end{aligned}$$

QED

Proof of Proposition 7. For $j \neq k$, the claims are immediate from equation (27) and propositions 5 and 6.

For $j = k$, there are two cases. If $\beta_{k-1}^{(k)} > 0$, then $\hat{y}_{k-1}^{(k)} > 0$ and $\hat{p}_{k-1}^{(k)} < 0$. Since both terms on the right in equation (A15) are positive, it follows that $\hat{p}_k^{(k)} < \hat{p}_{k-1}^{(k)} < 0$. This argument always holds if $\rho \leq 1$, and it holds for $\rho > 1$ if $\beta_{k-1}^{(k)} > 0$.

If $\rho > 1$ and $\beta_{k-1}^{(k)} < 0$, then $\hat{y}_j^{(k)} < 0$ and $\hat{p}_j^{(k)} > 0$, for all $j < k$. In addition, since $\beta_k^{(k)} > 0$, in this case $\hat{y}_j^{(k)} < 0$ and $\hat{p}_j^{(k)} > 0$, for all $j > k$. Since $\Sigma_{j=1}^{J}\nu_j\hat{p}_j^k = 0$, it follows that $\hat{p}_k^{(k)} < 0$. QED

Proof of Proposition 8. For $h \notin (b_{k-1}, b_k)$, the claim is immediate from propositions 4 and 7. For skill bin k, note that $\hat{w}(b_{k-1}) = \hat{p}_{k-1} > 0$, and $\hat{w}(h)$ is increasing in h for $h \in (b_{k-1}, b_k)$. QED

A3. An Example with Wage Declines

For an example where wages fall for some workers, let $J = 3$ and $k = 2$, and let the skill distribution be discrete, also with three types. Let h_i, ℓ_j, and $j = 1, 2, 3$ be the skill types and the number of workers of each type. The parameters are

$$\begin{aligned}
&x_3 = 10{,}000, && x_2 = 4, && x_1 = 1, && x_2' = 1.01x_2,\\
&h_3 = 10{,}000, && h_2 = 4, && h_1 = 0.95, &&\\
&\gamma_3 = 0.99, && \gamma_2 = 0.0090, && \gamma_1 = 0.0010, &&\\
&\ell_3 = 0.988912, && \ell_2 = 0.007991, && \ell_1 = 0.003097, &&\\
&\eta = 0.22, && \omega = 0.5, \quad \text{and} && \rho = 1.2. &&
\end{aligned}$$

The vast majority of firms have technology x_3, the vast majority of the workforce has skill $h_3 = x_3$, and these levels are much higher than the others. Hence, the increase in technology x_2 leaves final output virtually unchanged, and the price change at x_1 firms depends almost entirely on their own output change. In the initial equilibrium, all workers with skill h_3 are employed at firms with technology x_3, and all with skill h_2 are matched with technology x_2. Workers with skill h_1 are divided between firms with technologies x_1 and x_2. The increase in x_2 reallocates some additional h_1 workers to x_1 firms, and p_1 falls. Workers with skill h_1 take a wage cut equal to the decline in p_1.

A4. Skill-Biased Technical Change

Here we show that for each k, both of the terms on the right in equation (28) have order ε. For the first term, note that by construction

$$\begin{aligned}
\hat{\phi}_x(b_k, x_k) - \hat{\phi}_x(b_k, x_{k+1}) &= \hat{\phi}_x(b_k, x_k) - \hat{\phi}_x(b_k, x_k + \varepsilon)\\
&\approx \frac{\phi_x}{\phi} - \frac{\phi_x + \varepsilon\phi_{xx}}{\phi + \varepsilon\phi_x}\\
&\approx \varepsilon\frac{\phi_x}{\phi}\left(\frac{\phi_x}{\phi} - \frac{\phi_{xx}}{\phi_x}\right),
\end{aligned}$$

where ϕ and its derivatives are evaluated at (b_k, x_k).

For the second term, first note that

$$\begin{aligned}
\hat{\Psi}_k - \hat{\Psi}_{k+1} &= \frac{\int_{b_{k-1}}^{b_k}\phi_x(h, x_k)g(h)\,dh}{\int_{b_{k-1}}^{b_k}\phi(h, x_k)g(h)\,dh} - \frac{\int_{b_k}^{b_{k+1}}\phi_x(h, x_k + \varepsilon)g(h)\,dh}{\int_{b_k}^{b_{k+1}}\phi(h, x_k + \varepsilon)g(h)\,dh}\\
&\approx \frac{\phi_x(\bar{h}_k, x_k)}{\phi(\bar{h}_k, x_k)} - \frac{\phi_x(\bar{h}_{k+1}, x_k) + \varepsilon\phi_{xx}(\bar{h}_{k+1}, x_k)}{\phi(\bar{h}_{k+1}, x_k) + \varepsilon\phi_x(\bar{h}_{k+1}, x_k)},
\end{aligned}$$

where

$$\bar{h}_k \equiv \int_{b_{k-1}}^{b_k} hg(h)\,dh, \qquad \text{all } k,$$

is the average value in skill bin k. To approximate $\bar{h}_{k+1}$ in terms of $\bar{h}_k$, let $H(x)$ denote the inverse matching function in the continuous framework: technology x is paired with skill $H(x)$. Then, by construction, $\bar{h}_k \approx H(x_k)$ for all k, so

$$\bar{h}_{k+1} \approx \bar{h}_k + \varepsilon H'(x_k), \qquad \text{all } k.$$

Hence,

$$\begin{aligned}
\hat{\Psi}_k - \hat{\Psi}_{k+1} &\approx \frac{\phi_x}{\phi} - \frac{\phi_x + \varepsilon(\phi_{xh}H' + \phi_{xx})}{\phi + \varepsilon(\phi_h H' + \phi_x)} \\
&\approx \varepsilon \frac{\phi_x(\phi_h H' + \phi_x) - \phi(\phi_{xh}H' + \phi_{xx})}{\phi[\phi + \varepsilon(\phi_h H' + \phi_x)]} \\
&\approx \varepsilon \frac{\phi_x}{\phi}\left(\frac{\phi_h H' + \phi_x}{\phi} - \frac{\phi_{xh}H' + \phi_{xx}}{\phi_x}\right),
\end{aligned}$$

so this term also has order ε.

A5. *The Gain from PAM*

For the approximation in equation (29), note that $F(\bar{y}, \ldots, \bar{y}) = \bar{y}$, and

$$\begin{aligned}
F_j &= \left(\sum_k \gamma_k y_k^{(\rho-1)/\rho}\right)^{1/(\rho-1)} \gamma_j y_j^{-1/\rho}\Bigg|_{\bar{y}} = \gamma_j, \qquad \text{all } j, \\
F_{ji} &= \frac{1}{\rho}\bar{y}^{(2-\rho)/\rho}\gamma_j y_j^{-1/\rho}\gamma_i y_i^{-1/\rho}\Bigg|_{\bar{y}} = \frac{1}{\rho}\bar{y}^{-1}\gamma_j\gamma_i, \qquad \text{all } i \neq j, \\
F_{jj} &= \frac{1}{\rho}\bar{y}^{-1}\gamma_j^2 - \frac{1}{\rho}\bar{y}^{1/\rho}\gamma_j y_j^{-1-1/\rho}\Bigg|_{\bar{y}} \\
&= \frac{1}{\rho}\bar{y}^{-1}\gamma_j^2 - \frac{1}{\rho}\bar{y}^{-1}\gamma_j, \qquad \text{all } j.
\end{aligned}$$

Hence,

$$\sum_j F_j(y_j - \bar{y}) = 0,$$

$$\sum_j \sum_i F_{ji}(y_j - \bar{y})(y_i - \bar{y}) = -\frac{1}{\rho}\frac{\sigma_y^2}{\bar{y}}.$$

For the approximation in equation (31), note that $Z(\bar{x}, \ldots, \bar{x}) = q(\bar{x})^{1-\rho}$ and $Z_i = z_0\gamma_i$ for all i, where $z_0 > 0$ is a constant. Hence, $\Sigma_i Z_i(x_i - \bar{x}) = 0$, and $y_{j\mathrm{NP}}$ is as in equation (30).

For the approximations to q and q', use Euler's theorem to find that

$$q(\bar{x}) \approx \bar{x}\phi\left[1 + \left(\frac{\bar{h}}{\bar{x}}\right)^2 \frac{\phi_{hh}}{\phi}\frac{c_h^2}{2}\right],$$

$$q'(\bar{x}) \approx \phi_x\left[1 + \left(\frac{\bar{h}}{\bar{x}}\right)^2 \frac{\phi_{hhx}}{\phi_x}\frac{c_h^2}{2}\right],$$

where ϕ and its derivatives are evaluated at $(\bar{h}/\bar{x}, 1)$. Under the alignment assumption, $\bar{h}/\bar{x} = a_H$. Then by straightforward calculation,

$$\frac{a_H^2 \phi_{hh}}{\phi} = -A,$$
$$\frac{a_H^2 \phi_{hhx}}{\phi_x} = -BA,$$

where A and B are as in expression (34) and $\rho\phi_x/\phi = 1$. Hence, $\rho\bar{x}q'/q$ is as in expression (36).

A6. Wage Effects of Heterogeneous Technologies

For economies that satisfy the alignment assumption, the wage change from eliminating heterogeneity in technologies is

$$\begin{aligned}\Delta \ln w(\bar{h}_j) &\approx \chi_0 - \frac{1}{\rho}\left(\ln x^H - \ln x_j\right) + \left(\ln \phi\left(a_H, \frac{x^H}{x_j}\right) - \ln \phi(a_H, 1)\right),\\ &\approx \chi_0 - \frac{1}{\rho}\Delta_{xj} + \varepsilon_x \Delta_{xj} + \frac{1}{2}\varepsilon_{xx}\Delta_{xj}^2,\end{aligned}$$

where

$$\varepsilon_x \equiv \frac{\partial \ln \phi(h, x)}{\partial \ln x},$$
$$\varepsilon_{xx} \equiv \frac{\partial}{\partial \ln x}\left(\frac{x\phi_x(h, x)}{\phi(h, x)}\right).$$

For the elasticities, note that

$$\begin{aligned}\phi(h, x) &\equiv \left[\omega h^{(\eta-1)/\eta} + (1-\omega)x^{(\eta-1)/\eta}\right]^{\eta/(\eta-1)},\\ \phi_x(h, x) &= (1-\omega)x^{-1/\eta}\phi(h, x)^{1/\eta},\\ \phi_{xx}(h, x) &= \frac{1}{\eta}(1-\omega)x^{-1/\eta}\phi(h, x)^{1/\eta}\left(\frac{\phi_x}{\phi} - x^{-1}\right),\end{aligned}$$

so

$$\begin{aligned}\frac{x\phi_x}{\phi} &= (1-\omega)\phi\left(\frac{h}{x}, 1\right)^{-(\eta-1)/\eta},\\ \frac{x^2\phi_{xx}}{\phi} &= \frac{1}{\eta}(1-\omega)\phi\left(\frac{h}{x}, 1\right)^{-(\eta-1)/\eta}\left(\frac{x\phi_x}{\phi} - 1\right).\end{aligned}$$

Note, too, that

$$\phi(a_H, 1)^{-(\eta-1)/\eta} = \frac{1}{\rho(1-\omega)}.$$

Hence, evaluating the elasticities at $(h, x) = (a_H, 1)$ gives

$$\begin{aligned}
\varepsilon_x &= \frac{x\phi_x}{\phi} = (1-\omega)\phi(a_H,1)^{-(\eta-1)/\eta} = \frac{1}{\rho},\\
\varepsilon_{xx} &= x\left[\frac{\phi_x}{\phi} - x\left(\frac{\phi_x}{\phi}\right)^2 + \frac{x\phi_{xx}}{\phi}\right],\\
&= \frac{1}{\rho} - \frac{1}{\rho^2} + \frac{1}{\eta}(1-\omega)\phi(a_H,1)^{-(\eta-1)/\eta}\left(\frac{1}{\rho} - 1\right),\\
&= \frac{\rho-1}{\rho^2}\frac{\eta-1}{\eta}.
\end{aligned}$$

References

Acemoglu, Daron. 2002. "Technical Change, Inequality, and the Labor Market." *J. Econ. Literature* 40 (1): 7–72.

Acemoglu, Daron, and David Autor. 2011. "Skills, Tasks and Technologies: Implications for Employment and Earnings." In *Handbook of Labor Economics*, vol. 4B, edited by Orley Ashenfelter and David Card, 1043–171. Amsterdam: Elsevier.

Autor, David H., and David Dorn. 2013. "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market." *A.E.R.* 103 (5): 1553–97.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney. 2006. "The Polarization of the U.S. Labor Market." *A.E.R.* 96 (2): 189–94.

Autor, David H., Frank Levy, and Richard J. Murnane. 2003. "The Skill Content of Recent Technological Change: An Empirical Exploration." *Q.J.E.* 118 (4): 1279–333.

Becker, Gary S. 1973. "A Theory of Marriage: Part I." *J.P.E.* 81 (4): 813–46.

Costinot, Arnaud. 2009. "An Elementary Theory of Comparative Advantage." *Econometrica* 77 (4): 1165–92.

Costinot, Arnaud, and Jonathan Vogel. 2010. "Matching and Inequality in the World Economy." *J.P.E.* 118 (4): 747–86.

Goldin, Claudia, and Lawrence F. Katz. 2007. "Long-Run Changes in the Wage Structure: Narrowing, Widening, Polarizing." *Brookings Papers on Economic Activity* 2007 (2): 135–65.

———. 2008. *The Race between Education and Technology*. Cambridge, MA: Harvard Univ. Press.

Goldin, Claudia, and Robert Margo. 1992. "The Great Compression: the Wage Structure in the United States at Mid-Century." *Q.J.E.* 107 (1): 1–34.

Huang, Y., and W. F. McColl. 1997. "Analytical Inversion of General Tridiagonal Matrices." *J. Physics A* 30: 7919–33.

Katz, Lawrence F., and Kevin M. Murphy. 1992. "Changes in Relative Wages, 1963–1987: Supply and Demand Factors." *Q.J.E.* 107 (1): 35–78.

Krusell, Per, Lee E. Ohanian, Jose-Victor Rios-Rull, and Giovanni L. Violante. 2000. "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis." *Econometrica* 68 (5): 1029–53.

Lise, Jeremy, Costas Meghir, and Jean-Marc Robin. 2016. "Matching, Sorting and Wages." *Rev. Econ. Dynamics* 19:63–87.

Machin, Stephen, and John Van Reenen. 2007. "Changes in Wage Inequality." Special Paper no. 18, Centre for Economic Performance, London.

Neal, Derek, and Sherwin Rosen. 2000. "Theories of the Distribution of Earnings." In *Handbook of Income Distribution*, edited by Anthony B. Atkinson and François Bourguignon, 379–427. Amsterdam: Elsevier Science.

Ober, Harry. 1948. "Occupational Wage Differentials, 1907–1947." *Monthly Labor Rev.* 67 (2): 127–34.

Roy, A. D. 1950. "The Distribution of Earnings and Individual Output." *Econ. J.* 60 (September): 489–505.

Sattinger, Michael. 1975. "Comparative Advantage and the Distributions of Earnings and Abilities." *Econometrica* 43 (3): 455–68.

———. 1993. "Assignment Models and the Distribution of Earnings." *J. Econ. Literature* 31 (2): 831–80.

Song, Jae, David J. Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter. 2018. "Firming Up Inequality." Working Paper no. 750 (April), Fed. Reserve Bank Minneapolis.