Technology, Skill and the Wage Structure

Nancy L. Stokey*

University of Chicago and NBER

June 14, 2017 (Revised)

Abstract

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 to analyze the general equilibrium effects of technical change for a limited set of tasks. Output increases and price falls for tasks that are directly affected. The effects on employment depend on the elasticity of substitution across tasks. For high elasticities, employment expands to a group of more skilled workers. Hence for tasks farther up the technology ladder, employment falls, output declines, and prices and wages rise. For low elasticities, employment at affected tasks contracts among less skilled workers, as they shift to complementary tasks with unchanged technologies. In all cases, the output, price and wage changes are damped for more distant tasks, both above and below the affected group.
JEL Classification: D50, E24, O33, O40
KEYWORDS: technical change, human capital, wages

Department of Economics
University of Chicago
1126 E. 59th Street
Chicago, IL 60637
nstokey@uchicago.edu

1. 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. As documented by Goldin and Katz (2007), and described both succinctly and accurately in their title, the wage structure in the U.S. has seen "narrowing, widening, and polarizing" over the last 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

[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 1940's, followed by a slow widening in the 1950's and 1960's that accelerated in the 1970's and early 1980's; Katz and Murphy (1992) show that the college premium rose slowly in the 1960's, fell during the 1970's and rose sharply in the 1980's; and Autor, Katz and Kearney (2006), Machin and Van Reenen (2007), and Autor and Dorn (2013) document labor market polarization after 1980.

1

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 employment 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

3

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 will be shown below, the equilibrium can readily be re-interpreted 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 2 discusses the related literature. Section 3 presents the basic model and characterizes the competitive equilib-

rium. The main results are contained in section 4, 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 5 shows how the model can be used to address policy questions: the effects of a minimum wage and of opening to international trade or immigration. It also analyzes the effects of eliminating positively assortative matching and of eliminating technological heterogeneity. Section 6 concludes. Mathematical arguments and proofs are gathered in the Appendix.

2. 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 input 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 U.S.

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

5

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—affects 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 CV the production function is only required to be (strictly) log-supermodular. Thus the function here satisfies the requirement in CV, 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).

CV's 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 CV is extreme-biased if there exists a threshold technology with that property that the relevant distribution functions satisfy MLRP above the thresh-

old and the reverse condition below the threshold. Again, extreme biased shocks can be constructed as weighted sums of simple shocks.

Second, the results in CV are only about *relative* wage effects, while the model here delivers conclusions about wage levels, as well output and task price levels. The relationship between the shocks here and those in CV are discussed in more detail in section 5.

CV 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 will be 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.

3. 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 $j = 1, ..., J$, ordered so $0 < x_1 < x_2 < ... < x_J$. 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 (CRS) 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 will 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

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

where $\widetilde{y}_{ji}$ 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 demand for each task is

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

where $\{p_j, y_j\}_{j=1}^{J}$ and the aggregates y_F, 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 using 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 $H \equiv (h_{\min}, h_{\max})$, with $0 < h_{\min} < h_{\max} \leq \infty$, 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

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

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 $w(h)$, $h \in H$, 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 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 positively assortative matching: workers with higher skill h work with higher technologies x_j (Costinot, 2009). Consequently the equilibrium labor allocation is characterized by thresholds $h_{\min} = b_0 < b_1 < ... < b_{J-1} < b_J = h_{\max}$, where technology x_j employs workers with skill $h \in (b_{j-1}, b_j)$. We will 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

---

[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.

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:

a. income maximization by all types of labor,

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

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

c. labor market clearing,

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

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]

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

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

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

---

[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.

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

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

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

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 $\rho > 1$ (if $\rho < 1$), total revenue is increasing in j (decreasing in j).

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

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

$$j = 1, ..., J-1.$$

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 (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 'total productivity' of labor in the jth skill bin,

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

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

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

and write (10) in the more symmetric form

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

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, Figure 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,

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

Figure 1 displays potential wages as a function of h, for $J = 4$ 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, 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 (5). Hence the wage function $w(h)$ is defined by the upper envelop of the four curves, and the crossing points along the upper envelop are the thresholds b_j that divide the skill range into bins.

The four small circles 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 and from x_2 to x_3. But it falls moving from x_3 to x_4, so that worker chooses x_3.

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.

4. 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 $w^p(\cdot, x_k)$ 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, $j \neq k$. These price changes lower $w^p(\cdot, x_k)$ partway back toward its original level and raise all the other curves, $w^p(\cdot, x_j)$, $j \neq k$. 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 will 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,

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

where the weights

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

with $\sum_{j=1}^{J} \nu_j = 1$, 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

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

and the weighted average of the price changes is $\sum_{j=1}^{J} \nu_j \hat{p}_j = \hat{p}_F = 0$.

Consider first the short run effects, with labor immobile. Recall the definition of Ψ_j, all j, in (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,

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

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

$$\hat{y}_F^{SR} = \nu_k \hat{\Psi}_k > 0.$$

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 $\left\{b_j^{(k)}(\varepsilon)\right\}_{j=1}^{J-1}$ denote the solution to (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

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

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

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

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.

16

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 in the short run, $\hat{y}_F = \hat{y}_F^{SR}$.

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 (13) and use (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,

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

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

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

17

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

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 (23) to compare successive rows $j+1 > j \geq k$ in (20). Similarly, use the second line in (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 and above the kth move in the same direction, and all those at and below the $(k-1)^{\text{th}}$ 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,
  i. if $\rho = 1$, then $A_{k-1}^{(k)} > 0$ and $A_k^{(k)} > 0$;
  ii. if $\rho > 1$, then $A_k^{(k)} > 0$ and $A_{k-1}^{(k)}$ can have either sign; and
  iii. 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 (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 $\rho < 1$, 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 $\rho > 1$, 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 (22) and Lemma 2 that all thresholds shift upward. But even if one term in (22) is negative, 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:

for $\rho = 1$,

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

for $\rho > 1$,

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

and for $\rho < 1$,

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

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 $(k-1)^{\text{th}}$ can shift either way. For

19

$\rho < 1$, the thresholds at and below the $(k-1)^{\text{th}}$ shift upward, while those at and above the kth can shift either way.

C. Task outputs

From (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,

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

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.

20

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 (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 $j > k$,

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

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

21

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

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

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

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 is the same as the price change, and is 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 $\hat{w}(h) > 0$, 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 $\beta_k^{(k)} < 0$, 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 $\beta_{k-1}^{(k)} > 0$, 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.

22

5. EXAMPLES AND APPLICATIONS

In this section we will first look at several examples that illustrate the relationship between the results above and those in Costinot and Vogel (2010). The model will then be used to study several substantive questions—the effects of minimum wage legislation, of opening up to international trade, of positively assortative matching (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 will call such shifts *simple*.

In CV (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 $\gamma^o(\cdot)$ on X. A technology shift changes the density, from γ^o to γ^n.

CV study two types of shifts. A technology shift is *skill biased* if the densities satisfy the monotone likelihood ratio property (MLRP). By definition, this property holds if and only if

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

That is, the ratio of the new density to the old must be (weakly) increasing in x. Lemma 5 in CV 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 CV is *extreme biased* if there exists a threshold technology $\hat{x}$ with the property that the densities satisfy MLRP above $\hat{x}$ and satisfy the reverse property below $\hat{x}$. Lemma 6 in CV shows that after such a shift, there exists a skill

23

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.

i. 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 $\gamma^n(x)/\gamma^o(x)$ is unity except near x_k, where it shrinks almost zero, and near $x_k + \varepsilon$, where it explodes. Hence no shift of this type satisfies CV's definition of skill biased. The simple shifts considered in section 4 satisfy first order stochastic dominance (FOSD), but not MLRP.

ii. Skilled 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
$$\frac{D\gamma^o(x)}{\gamma^o(x)} > \frac{D\gamma^o(x')}{\gamma^o(x')}, \qquad \text{all } x < x',$$
where $D\gamma^o \equiv d\gamma^o/dx$. That is $D\gamma^o/\gamma^o$ must be a decreasing function. Suppose this is the case, so Lemma 5 in CV 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, and with $x_1 = a + \varepsilon/2$

and $x_J = b - \varepsilon/2$. Define the probabilities $\{\gamma_j\}_{j=1}^J$ by

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

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

$$\hat\gamma_{j+1} = \gamma_j, \qquad j = 1, ..., J.$$

Moreover, this shift clearly is isomorphic to the sum of J simple shifts of the type described in section 4.

Therefore, summing the changes in (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

$$A_k^{(k)} + A_k^{(k+1)} = \rho \left[\hat\phi_x(b_k, x_k) - \hat\phi_x(b_k, x_{k+1}) \right] - \left[\hat\Psi_k - \hat\Psi_{k+1} \right], \qquad \text{all } k. \qquad (28)$$

As shown in the Appendix, for each k, both of the terms on the right in (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 define in (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.

iii. 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 CV. 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. Minimum wage legislation

The model here can be used to assess the effects of minimum wage legislation. The analysis proceeds by truncating a small segment of workers at the bottom of the skill distribution, analyzing the effects on all wages, and then backing out the minimum wage that would prompt the change. The method is exactly as in section 4, except that the exogenous shock is a change in the supply of labor of a particular type, instead of a technology shock.

Fix the baseline economy and its competitive equilibrium, choose $\varepsilon > 0$, and suppose all individuals with $h < h_{\min} + \varepsilon$ are prohibited from working. 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,

$$\beta_0^{(w)} \equiv g(b_0),$$

$$\beta_j^{(w)} \equiv g(b_j)b_j^{(w)\prime}, \qquad j = 1, ..., J-1,$$

$$\beta_J^{(w)} \equiv 0.$$

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

$$A_1^{(w)} \equiv \frac{1}{\Psi_1}\phi(b_0, x_1)\beta_0^{(w)} > 0,$$

$$A_j^{(w)} \equiv 0, \qquad j = 2, ..., J-1.$$

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

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

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

Using (12) and the argument in the proof of Proposition 5, the effects on task outputs are

$$\hat{y}_j^{(w)} = -\frac{1}{\Psi_j}\left[d_j\frac{\psi_{j+1}}{\psi_j} + 1\right]\phi(b_{j-1}, x_j)\beta_{j-1}^{(w)} < 0, \qquad j = 1, ..., J-1,$$

$$\hat{y}_J^{(w)} = -\frac{1}{\Psi_J}\phi(b_{J-1}, x_J)\beta_{J-1}^{(w)} < 0,$$

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,

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

so the changes are damped farther up the task ladder,

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

Using the argument in the proof of Proposition 7, the proportionate price changes satisfy

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

Since $\sum_1^J \nu_j \hat{p}_j = 0$, prices rise for a set of tasks at the bottom of the ladder, and fall for the complementary set. That is, there exists $1 < k \leq J$ such that $\hat{p}_j > 0$ for $j < k$, and $\hat{p}_j \leq 0$ for $j \geq k$. 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, wages rise for workers in skill bins $1 \leq j < k$, 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.

Finally, notice that the new equilibrium is also the one that prevails if the baseline economy adopts a minimum wage of

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

Thus, minimum wage legislation hurts workers at the very bottom of the ladder, who lose their jobs, benefits workers in a range 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 near the bottom of the skill ladder who keep their jobs are the biggest winners.

Figure 2 displays, for a numerical example. The substitution elasticity between technology and skill is set at $\eta = 0.5$, and the two inputs are given 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 parameter $\lambda_F = 1.0$ and location $x_{\min} = 1$. The skill distribution is a truncated lognormal, with mean and variance of unity. In

summary, the parameters are

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

In each economy the minimum wage is chosen so that it puts 1% of the labor force out of work. The wage function in these economies is steep at the bottom end, so the wage at the first percentile is in each case about 22% higher than at the bottom. Not surprisingly, in each case the required minimum wage is about 22% higher than the lowest wage in the baseline economy. Because skill and wage are so low at the bottom of the skill distribution, the reduction in final output is between 0.22% and 0.30%, with lower elasticities producing larger losses. Virtually all of the loss is born by those who lose their jobs: the total wage bill for those who remain employed is almost unchanged. But in line with the theory, Figure 2 shows that wages rise slightly for workers who remain employed in the lowest skill bins, and decline slightly for workers farther up the skill ladder.

C. International trade, immigration

The effects of international trade, for two special cases, can be analyzed in a similar way. We will assume that all tasks are costlessly tradeable, and that both countries produce the same set of tasks and have the same production function for final output.

Consider a large country whose workers have skills in the range $[h_{\min} + \Delta, h_{\max}]$. Suppose this country opens up to trade with a small, less-skilled partner, one with workers who have skills in the range $[h_{\min}, h_{\min} + \Delta]$. Before opening, the large country's economy is like the one in the previous section, with the minimum wage. After opening, the integrated world economy is like the one in the previous section, without a minimum wage. Hence the effects on the thresholds, task outputs, prices and wages for the large country are exactly the reverse of those resulting from a minimum wage.

29

With trade, all bin thresholds in the large country shift downward, so all workers experience task upgrading. Output of every task is higher in the integrated economy, and the proportionate output increases are damped farther up the task ladder. Hence the proportionate price changes satisfy

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

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 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, skill-poor country. In the large country, which has more heterogeneity across workers, there are losers as well as winners.

Figure 2 illustrates the wage effects for a large, skill-rich country, one with 99% of the world population, all of whom are at the top of the world skill distribution, that opens up to trade with a small, poorly-skilled partner, one with 1% of the world population, all of whom are at the bottom of the skill distribution. For the skill-rich country, opening to trade results in slight wage reductions in the lowest skill bins and slight wage gains at the top of the skill distribution, exactly the reverse of the changes in Figure 2.

The effects of trade for a large, less-skilled country that opens to trade with a small, skill-rich partner can be analyzed analytically in exactly the same way. The

effects of opening trade between two countries of similar size or with overlapping skill distributions is more complex, but could easily be studied numerically.

Finally, in all cases immigration produces exactly the same effects for workers in the host country.

D. Gain from positively assortative matching (PAM)

To analyze the gains from positively assortative matching (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 firms and their sizes are indeterminate, but also irrelevant.

If skill is observable, the equilibrium is exactly as before, and the economy-wide average wage is $\mathrm{E}[w(h)] = y_F$. If skill is unobservable, 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 will 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 $(y_1, ..., y_J)$ of task inputs, and let $\overline{y} = \sum \gamma_j y_j$, $\sigma_y^2 = \sum \gamma_j (y_j - \overline{y})^2$, and $c_y \equiv \sigma_y / \overline{y}$ denote the mean, variance, and coefficient

of variation (CoV). Then final output is

$$F(y_1, ..., y_J) \equiv \left(\sum_j \gamma_j y_j^{(\rho-1)/\rho} \right)^{\rho/(\rho-1)}$$

$$\approx \bar{y} \left(1 - \frac{1}{\rho} \frac{1}{2} c_y^2 \right), \tag{29}$$

where the second line approximates around $(\bar{y}, ..., \bar{y})$. 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 will be 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 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,

$$q(x_j) \equiv \mathrm{E}_h \left[\phi(h, x_j) \right], \qquad \text{all } j.$$

Task outputs in the NP economy are

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

where

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

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

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

32

so task outputs have mean, variance, and CoV

$$\overline{y}_{NP} \approx q(\overline{x}), \qquad \sigma^2_{yNP} \approx (\rho q')^2 \sigma^2_x, \qquad c_{yNP} \approx \frac{\rho \overline{x} q'}{q} c_x. \tag{31}$$

To analyze the equilibrium with PAM, a tractable family of economies are 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 $\sum_{i=1}^{j} \gamma_i \approx G(a_H x_j)$, all x_j.

Under the Alignment Assumption, skill has approximately the same distribution as technology, but scaled by a_H. Thus, $\overline{h}/\overline{x} = a_H$, 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

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

with mean, variance and CoV

$$\overline{y}_P \approx \overline{x}\phi, \qquad \sigma^2_{yP} \approx \phi^2 \sigma^2_x, \qquad c_{yP} \approx c, \tag{32}$$

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

$$\overline{y}_{NP} \approx \overline{x}\phi \left(1 - Ac^2/2 \right), \qquad c_{yNP} \approx \frac{1 - BAc^2/2}{1 - Ac^2/2} c, \tag{33}$$

---

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

where

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

$$B(\rho, \eta) \equiv (\rho - 1) - \frac{1}{\eta} (\rho - 2),$$

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

$$c^2/2 < 1/A, \qquad \text{and} \qquad Bc^2/2 < 1/A. \tag{35}$$

Recall from (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 (32) and (33),

$$\overline{y}_{NP} \approx \overline{x}\phi \left(1 - Ac^2/2\right) < \overline{x}\phi \approx \overline{y}_P.$$

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 (32) and (33), it reinforces or mitigates the mean effect,

$$\left(1 - \frac{1}{2}\frac{1}{\rho}c_{yP}^2\right) \gtreqless \left(1 - \frac{1}{2}\frac{1}{\rho}c_{yNP}^2\right),$$

or equivalently as $c_{yP} \lesseqgtr c_{yNP}$, or as

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

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

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

34

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 (35) and (37).

CASE A: As $\rho \downarrow 1$, (34) implies $A \to 0$ and $B \to 1/\eta$, and (36) implies $\lim_{\rho \to 1} c_{yNP} = c_{yP}$. 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 $\rho \to \infty$, (34) implies $A \to 0$ and $BA \to (\eta - 1)/\eta^2 < 0$, so (36) implies $\lim_{\rho \to \infty} c_{yNP} > c_{yP}$. 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 (36) that $\lim_{\eta \to 1} B = 1$, 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,

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

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 will 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 skill allocation, 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 $y_{jHT} = y_{HT} = y_F$, all j. Hence the required technology level x_{HT} satisfies

$$y_{HT} = \mathrm{E}_h \left[\phi(h, x_{HT}) \right].$$

Then from (2) and (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 (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 (38). Output can rise because the technology improves, because the average skill of co-workers 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 expand 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 $\phi(h_j, x_{HT})$. Hence the wage change is

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

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

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

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

The coefficient on the quadratic term in (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 (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 in the baseline economy. In these examples technology inequality *reduces* wage inequality, because of substantial price effects.

6. CONCLUSION

The analysis here has focussed on the effects of technology changes, but the framework could also be extended and used to examine other questions. As illustrated in the examples in section 5, it could be used to study the wage shifts arising from changes in minimum wage, trade, and immigration 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 micro-

foundation for the match surplus function, a function that frictional search models take as exogenous. As a consequence, the model here produces a non-degenerate 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 last 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. Large increases in wage inequality lead, 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.

---

[5]For example, see Song, et. al. (2016).

REFERENCES

[1] Acemoglu, Daron. 2002. Technical change, inequality, and the labor market, *Journal of Economic Literature*, 40(1): 7-72.

[2] Acemoglu, Daron, and David Autor. 2011. Skills, tasks and technologies: implications for employment and earnings. In *Handbook of Labor Economics, Vol. 4, Part B,* edited by Orley Ashenfelter and David Card, pp. 1043-1171. Amsterdam: Elsevier.

[3] Autor, David H. and David Dorn. 2013. The growth of low-skill service jobs and the polarization of the US labor market, *American Economic Review*, 103(5): 1553-1597.

[4] Autor, David H., Lawrence F. Katz, and Melissa S. Kearney. 2006. The polarization of the U.S. labor market, *American Economic Review: Papers and Proceedings*, 96(2): 189-194.

[5] Autor, David H., Frank Levy, and Richard J. Murnane. 2003. The skill content of recent technological change: an empirical exploration, *Quarterly Journal of Economics*, 118(4): 1279-1333.

[6] Becker, Gary S. 1973. A theory of marriage: part I, *Journal of Political Economy*, 81(4): 813-846.

[7] Costinot, Arnaud. 2009. An elementary theory of comparative advantage, *Econometrica* 77(4): 1165-92.

[8] Costinot, Arnaud, and Jonathan Vogel. 2010. Matching and inequality in the world economy, *Journal of Political Economy*, 118(4): 747-786.

[9] Goldin, Claudia, and Lawrence F. Katz. 2007.Long-run changes in teh wage structure: narrowing, widening, polarizing, *Brookings Papers on Economic Activity* 2: 135–165.

[10] Goldin, Claudia, and Lawrence F. Katz. 2008. *The Race between Education and Technology*, Harvard University Press.

[11] Goldin, Claudia and Robert Margo. 1992. The great compression: the wage structure in the United States at mid-century, *Quarterly Journal of Economics*, 107(1): 1-34.

[12] Huang, Y. and W. F. McColl. 1997. Analytical inversion of general tridiagonal matrices, *Journal of Physics A: Mathematical and General*, 30: 7919-33.

[13] Katz, Lawrence F. and Kevin M. Murphy. 1992. Changes in relative wages, 1963-1987: supply and demand factors, *Quarterly Journal of Economics*, 107(1): 35-78.

[14] 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): 1029-1053.

[15] Lise, Jeremy, Costas Meghir, and Jean-Marc Robin. 2016. Matching, sorting and wages, *Review of Economic Dynamics*, 19: 63-87.

[16] Machin, Stephen, and John Van Reenen. 2007. Changes in wage inequality, Center for Economic Performance Special Paper No. 18.

[17] Neal, Derek, and Sherwin Rosen. 2000. Theories of the distribution of earnings, Chapter 7. *Handbook of Income Distribution*, ed. by A.B. Atkinson and F. Bourguignon, Elsevier Science.

[18] Ober, Harry. 1948. Occupational wage differentials, 1907-1947, *Monthly Labor Review*, 67(2): 127-134.

[19] Roy, A.D. 1950. The distribution of earnings and individual output, *Economic Journal*, 60: 489-505.

[20] Sattinger, Michael. 1975. Comparative advantage and the distributions of earnings and abilities, *Econometrica* 43(3): 455-468.

[21] Sattinger, Michael. 1993. Assignment models and the distribution of earnings, *Journal of Economic Literature,* 31(2): 831-880.

[22] Song, Jae, David J. Price, Fatih Guvenen, and Nicholas Bloom, Till von Wachter, Firming up inequality, working paper, October 2016.

APPENDIX

A. Proof of Proposition 1

PROOF OF PROPOSITION 1: Use (19) in (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 (15) and the fact that $\beta_0^{(k)} = \beta_J^{(k)} = 0$, the second uses (5), and the third uses (9). From (13), the term in brackets in the last line is zero, for all j. ∎

B. Matrix M and proofs of Results 2 - 8

Differentiate (13) and use (11) to get

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

Write this in matrix form as

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

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

$$
\begin{aligned}
a_j &\equiv \rho\delta_j - (d_j + c_{j+1}) > 0, & j &= 1,...,J-1, & (40) \\
\delta_j &\equiv \left[\hat{\phi}_h(b_j, x_{j+1}) - \hat{\phi}_h(b_j, x_j)\right]/g(b_j) > 0, & j &= 1,...,J-1, \\
c_j &\equiv -\frac{1}{\Psi_j}\phi(b_{j-1}, x_j) < 0, & j &= 2,...,J, \\
d_j &\equiv -\frac{1}{\Psi_j}\phi(b_j, x_j) < 0, & j &= 1,...,J-1.
\end{aligned}
$$

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

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

$$
\begin{aligned}
\theta_0 &\equiv 1, & \theta_1 &\equiv a_1, & (41) \\
\theta_i &\equiv a_i\theta_{i-1} - c_i d_i\theta_{i-2}, & i &= 2,...,J-1; \\
\psi_J &\equiv 1, & \psi_{J-1} &\equiv a_{J-1}, & (42) \\
\psi_i &\equiv a_i\psi_{i+1} - c_{i+1}d_{i+1}\psi_{i+2}, & i &= J-2,...,1.
\end{aligned}
$$

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

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

$$
\begin{aligned}
\theta_{i-1} + c_i\theta_{i-2} &> 0, & i &= 2,...,J-1, & (43) \\
\psi_i + d_i\psi_{i+1} &> 0, & i &= J-2,...,1. & (44)
\end{aligned}
$$

PROOF OF LEMMA A1: Use (40) in (41) to find that

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

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

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

Since $\theta_1 = a_1 > 0$, and

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

by induction (43) holds. Similarly, use (40) in (42) to find that

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

so

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

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 (44) holds. ∎

PROOF OF LEMMA 2: The matrix M has elements (see Huang and McColl, 1997)

$$m_{nn} = \frac{1}{\theta_{J-1}}\theta_{n-1}\psi_{n+1}, \qquad n = 1, ..., J - 1, \tag{45}$$

$$m_{j+1,n} = -c_{j+1}\frac{\psi_{j+2}}{\psi_{j+1}}m_{j,n}, \qquad j = n, ..., J - 2,$$

$$m_{j-1,n} = -d_j\frac{\theta_{j-2}}{\theta_{j-1}}m_{j,n}, \qquad j = n, ..., 2.$$

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

$$q_{j+1} \equiv -c_{j+1}\frac{\psi_{j+2}}{\psi_{j+1}}, \qquad j = 1, ..., J - 2, \tag{46}$$

$$r_{j-1} \equiv -d_j\frac{\theta_{j-2}}{\theta_{j-1}}, \qquad j = 2, ..., J. \qquad ∎$$

PROOF OF LEMMA 3: From the definitions of $\hat{\phi}_x$ and $\hat{\Psi}_x$,

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

45

and since ϕ is a CES function,

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

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

$$\int_{b_{k-1}}^{b_k} \phi(h, x_k) \left[\rho \phi(b_k, x_k)^{1/\eta - 1} - \phi(h, x_k)^{1/\eta - 1} \right] g(h) dh \gtreqqless 0. \tag{47}$$

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

$$\int_{b_{k-1}}^{b_k} \phi(h, x_k) \left[\rho \phi(b_{k-1}, x_k)^{1/\eta - 1} - \phi(h, x_k)^{1/\eta - 1} \right] g(h) dh \lesseqqgtr 0. \tag{48}$$

Recall that $\phi(\cdot, x_k)$ is increasing in its first argument, and $\eta < 1$. For $\rho \geq 1$, the term in square brackets in (47) is positive over the range of integration, so $A_k^{(k)} > 0$. For $\rho \leq 1$, the term in square brackets in (48) is negative, so $A_{k-1}^{(k)} > 0$. In other cases the signs are ambiguous. ∎

PROOF OF PROPOSITION 4: For $\rho = 1$, the claims are immediate from (22) and Lemmas 2 and 3. For $\rho \neq 1$, the same is true for $k = 1$ and $k = J$, since (22) has only one term.

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

$$\begin{aligned}
\beta_k^{(k)} &= q_k m_{k-1, k-1} A_{k-1}^{(k)} + m_{k,k} A_k^{(k)} \\
&= \frac{\psi_{k+1}}{\theta_{J-1}} \left(-c_k \theta_{k-2} A_{k-1}^{(k)} + \theta_{k-1} A_k^{(k)} \right),
\end{aligned} \tag{49}$$

where the second line uses (45) and (46). Similarly, use the second line of (23), with $j = n = k$, in (22) to find that

$$\begin{aligned}
\beta_{k-1}^{(k)} &= m_{k-1, k-1} A_{k-1}^{(k)} + r_{k-1} m_{k,k} A_k^{(k)} \\
&= \frac{\theta_{k-2}}{\theta_{J-1}} \left(-\psi_k A_{k-1}^{(k)} - d_k \psi_{k+1} A_k^{(k)} \right).
\end{aligned} \tag{50}$$

Suppose $\rho > 1$. Then $A_k^{(k)} > 0$, so the second term in (49) 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

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

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

so $\left| A_{k-1}^{(k)} \right| < A_k^{(k)}$. Hence by Lemma A1 the sum in parenthesis in (49) is positive. In (50), the fact that $\left| A_{k-1}^{(k)} \right| < A_k^{(k)}$, 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 (50) 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 $\left| A_k^{(k)} \right| < A_{k-1}^{(k)}$. Hence by Lemma A1 the sum in parenthesis in (50) is positive. In (49), the fact that $\left| A_k^{(k)} \right| < A_{k-1}^{(k)}$, does not help in applying Lemma A1, so the sign is ambiguous. ∎

PROOF OF PROPOSITION 5: Recall from (40) that

$$
\phi(b_j, x_j)c_j = \phi(b_{j-1}, x_j)d_j, \qquad \text{all } j. \tag{51}
$$

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

$$
\begin{aligned}
\hat{y}_j^{(k)} &= \frac{1}{\Psi_j} \left[\phi(b_j, x_j)q_j - \phi(b_{j-1}, x_j) \right] \beta_{j-1}^{(k)} \\
&= \frac{1}{\Psi_j} \left[-c_j \frac{\psi_{j+1}}{\psi_j} \phi(b_j, x_j) - \phi(b_{j-1}, x_j) \right] \beta_{j-1}^{(k)} \\
&= -\frac{1}{\Psi_j} \left[d_j \frac{\psi_{j+1}}{\psi_j} + 1 \right] \phi(b_{j-1}, x_j) \beta_{j-1}^{(k)} \\
&\gtreqqless 0 \qquad \text{as} \qquad \beta_k^{(k)} \lesseqqgtr 0,
\end{aligned} \tag{52}
$$

where the second line uses the definition of q_j, the third uses (51), and the last uses Lemma A1 and Proposition 4. Similarly, for $j < k$, use the second line in (24) in (25), and the definition of r_{j-1} to find that

$$
\begin{aligned}
\hat{y}_j^{(k)} &= \frac{1}{\Psi_j} \left[\phi(b_j, x_j) - r_{j-1} \phi(b_{j-1}, x_j) \right] \beta_j^{(k)} \\
&= \frac{1}{\Psi_j} \left[\phi(b_j, x_j) + d_j \frac{\theta_{j-2}}{\theta_{j-1}} \phi(b_{j-1}, x_j) \right] \beta_j^{(k)} \\
&= \frac{1}{\Psi_j} \left[1 + c_j \frac{\theta_{j-2}}{\theta_{j-1}} \right] \phi(b_j, x_j) \beta_j^{(k)} \\
&\gtreqless 0 \qquad \text{as} \qquad \beta_{k-1}^{(k)} \gtreqless 0, \qquad j < k.
\end{aligned}
\tag{53}
$$

For $j = k$, the first term in (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

$$
\hat{p}_{k-1} - \hat{p}_k = \left[\hat{\phi}_h(b_{k-1}, x_k) - \hat{\phi}_h(b_{k-1}, x_{k-1}) \right] \frac{\beta_{k-1}^{(k)}}{g(b_{k-1})} + \hat{\phi}_x(b_{k-1}, x_k).
\tag{54}
$$

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 (26) are positive. If in addition $\beta_k^{(k)} \geq 0$, then 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

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

48

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 $\hat{y}_{k+1}^{(k)} > 0$. ∎

PROOF OF PROPOSITION 6: For $j > k$, use (52), 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 (53), 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

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

PROOF OF PROPOSITION 7: For $j \neq k$, the claims are immediate from (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 (54) are positive, it follows that $\hat{p}_k^{(k)} < \hat{p}_{k-1}^{(k)} < 0$. This argument always holds if $\rho \leq 1$, and 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 $\hat{p}_j^{(k)} > 0$, all $j < k$. In addition, since $\beta_k^{(k)} > 0$, in this case $\hat{y}_j^{(k)} < 0$ and $\hat{p}_j^{(k)} > 0$, all $j > k$. Since $\Sigma_{j=1}^J \nu_j \hat{p}_j^{(k)} = 0$, it follows that $\hat{p}_k^{(k)} < 0$. ∎

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 $h \in (b_{k-1}, b_k)$. ■

C. 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, $j = 1, 2, 3$, be the skill types and the number of workers of each type. The parameters are

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

The vast majority of firms have technology x_3, and 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 decline in p_1.

D. Skill-biased technical change

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

$$
\hat{\phi}_x(b_k, x_k) - \hat{\phi}_x(b_k, x_{k+1}) = \hat{\phi}_x(b_k, x_k) - \hat{\phi}_x(b_k, x_k + \varepsilon)
$$

$$\approx \quad \frac{\phi_x}{\phi} - \frac{\phi_x + \varepsilon \phi_{xx}}{\phi + \varepsilon \phi_x}$$

$$\approx \quad \varepsilon \frac{\phi_x}{\phi} \left(\frac{\phi_x}{\phi} - \frac{\phi_{xx}}{\phi_x} \right),$$

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

For the second term, first note that

$$\hat{\Psi}_k - \hat{\Psi}_{k+1} = \frac{\int_{b_{k-1}}^{b_k} \phi_x(h, x_k) g(h) dh}{\int_{b_{k-1}}^{b_k} \phi(h, x_k) g(h) dh} - \frac{\int_{b_k}^{b_{k+1}} \phi_x(h, x_k + \varepsilon) g(h) dh}{\int_{b_k}^{b_{k+1}} \phi(h, x_k + \varepsilon) g(h) dh}$$

$$\approx \quad \frac{\phi_x(\overline{h}_k, x_k)}{\phi(\overline{h}_k, x_k)} - \frac{\phi_x(\overline{h}_{k+1}, x_k) + \varepsilon \phi_{xx}(\overline{h}_{k+1}, x_k)}{\phi(\overline{h}_{k+1}, x_k) + \varepsilon \phi_x(\overline{h}_{k+1}, x_k)},$$

where

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

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

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

Hence

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

$$\approx \quad \varepsilon \frac{\phi_x \left(\phi_h H' + \phi_x \right) - \phi \left(\phi_{xh} H' + \phi_{xx} \right)}{\phi \left[\phi + \varepsilon \left(\phi_h H' + \phi_x \right) \right]}$$

$$\approx \quad \varepsilon \frac{\phi_x}{\phi} \left(\frac{\phi_h H' + \phi_x}{\phi} - \frac{\phi_{xh} H' + \phi_{xx}}{\phi_x} \right),$$

so this term also has order ε.

E. The gain from PAM

For the approximation in (29), note that $F(\overline{y}, ..., \overline{y}) = \overline{y}$, and

$$F_j = \left(\sum_k \gamma_k y_k^{(\rho-1)/\rho} \right)^{1/(\rho-1)} \gamma_j y_j^{-1/\rho} \Big|_{\overline{y}} = \gamma_j, \qquad \text{all } j,$$

$$F_{ji} = \frac{1}{\rho}\overline{y}^{(2-\rho)/\rho}\gamma_j y_j^{-1/\rho}\gamma_i y_i^{-1/\rho}|_{\overline{y}} \qquad\qquad = \frac{1}{\rho}\overline{y}^{-1}\gamma_j\gamma_i, \qquad \text{all } i \neq j,$$

$$F_{jj} = \frac{1}{\rho}\overline{y}^{-1}\gamma_j^2 - \frac{1}{\rho}\overline{y}^{1/\rho}\gamma_j y_j^{-1-1/\rho}|_{\overline{y}}$$

$$= \frac{1}{\rho}\overline{y}^{-1}\gamma_j^2 - \frac{1}{\rho}\overline{y}^{-1}\gamma_j, \qquad\qquad\qquad \text{all } j.$$

Hence

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

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

For the approximation in (31), note that $Z(\overline{x},...,\overline{x}) = q(\overline{x})^{1-\rho}$, and $Z_i = z_0\gamma_i$, all i, where $z_0 > 0$ is a constant. Hence $\sum_i Z_i\,(x_i - \overline{x}) = 0$, and y_{jNP} is as in (30).

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

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

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

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

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

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

F. Wage effects of heterogeneous technologies

For economies that satisfy the Alignment Assumption, the wage change from eliminating heterogeneity in technologies is

$$\Delta\ln w(\overline{h}_j) \approx \chi_0 - \frac{1}{\rho}\left[\ln x^H - \ln x_j\right] + \left[\ln\phi\left(a_H, x^H/x_j\right) - \ln\phi\left(a_H, 1\right)\right],$$

$$\approx \chi_0 - \frac{1}{\rho}\Delta_{xj} + \varepsilon_x\Delta_{xj} + \frac{1}{2}\varepsilon_{xx}\Delta_{xj}^2,$$

where

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

For the elasticities, note that

$$\phi(h, x) \equiv \left[\omega h^{(\eta-1)/\eta} + (1 - \omega) x^{(\eta-1)/\eta} \right]^{\eta/(\eta-1)}$$

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

$$\phi_{xx}(h, x) = \frac{1}{\eta} (1 - \omega) x^{-1/\eta} \phi(h, x)^{1/\eta} \left(\frac{\phi_x}{\phi} - x^{-1} \right),$$

so

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

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

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

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

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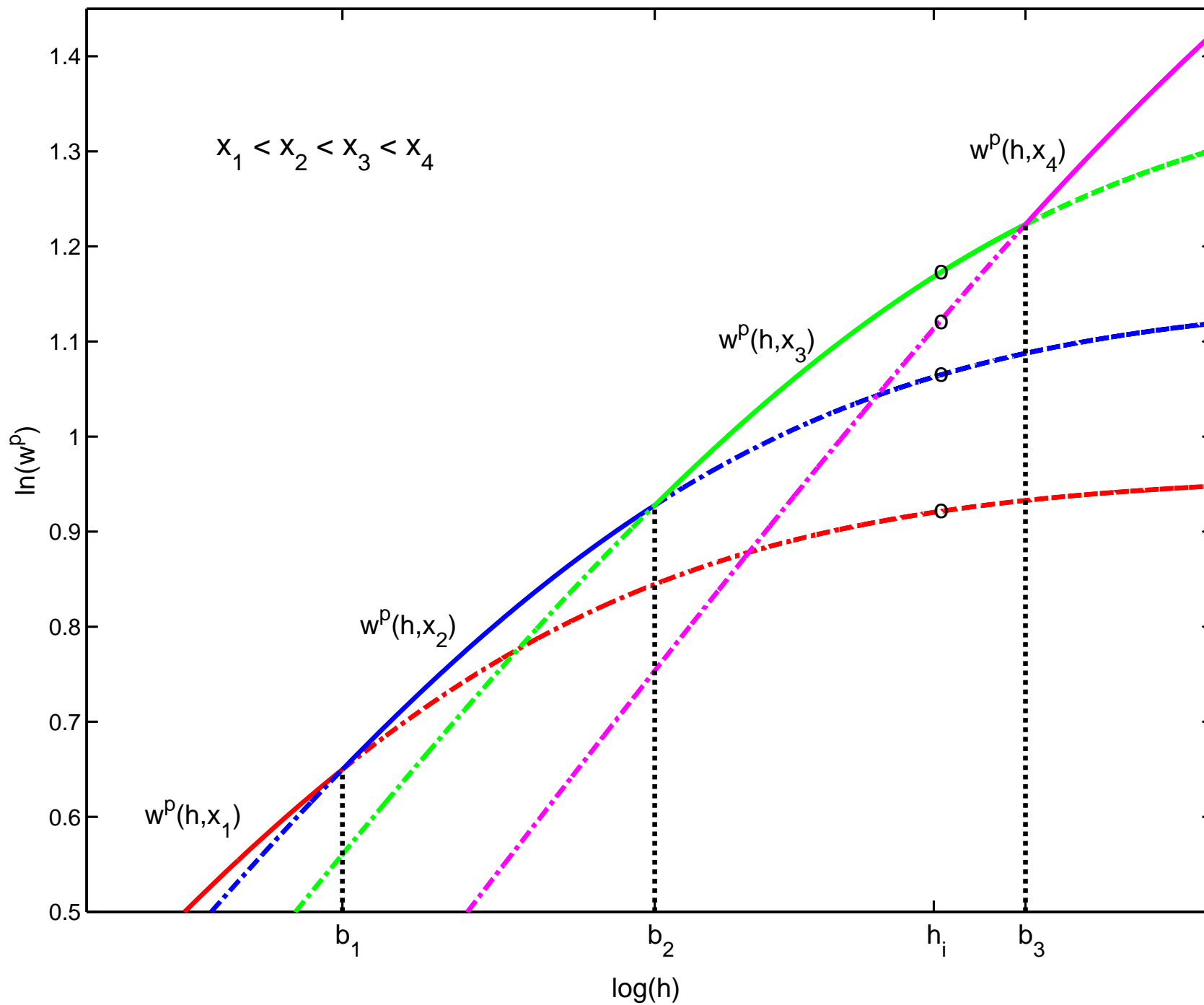\end{aligned}$$

53

Figure 1: potential wages

Figure 2: wage changes from minimum wage legislation

$\ln(w_{MW}) - \ln(w)$

$\rho = 6$
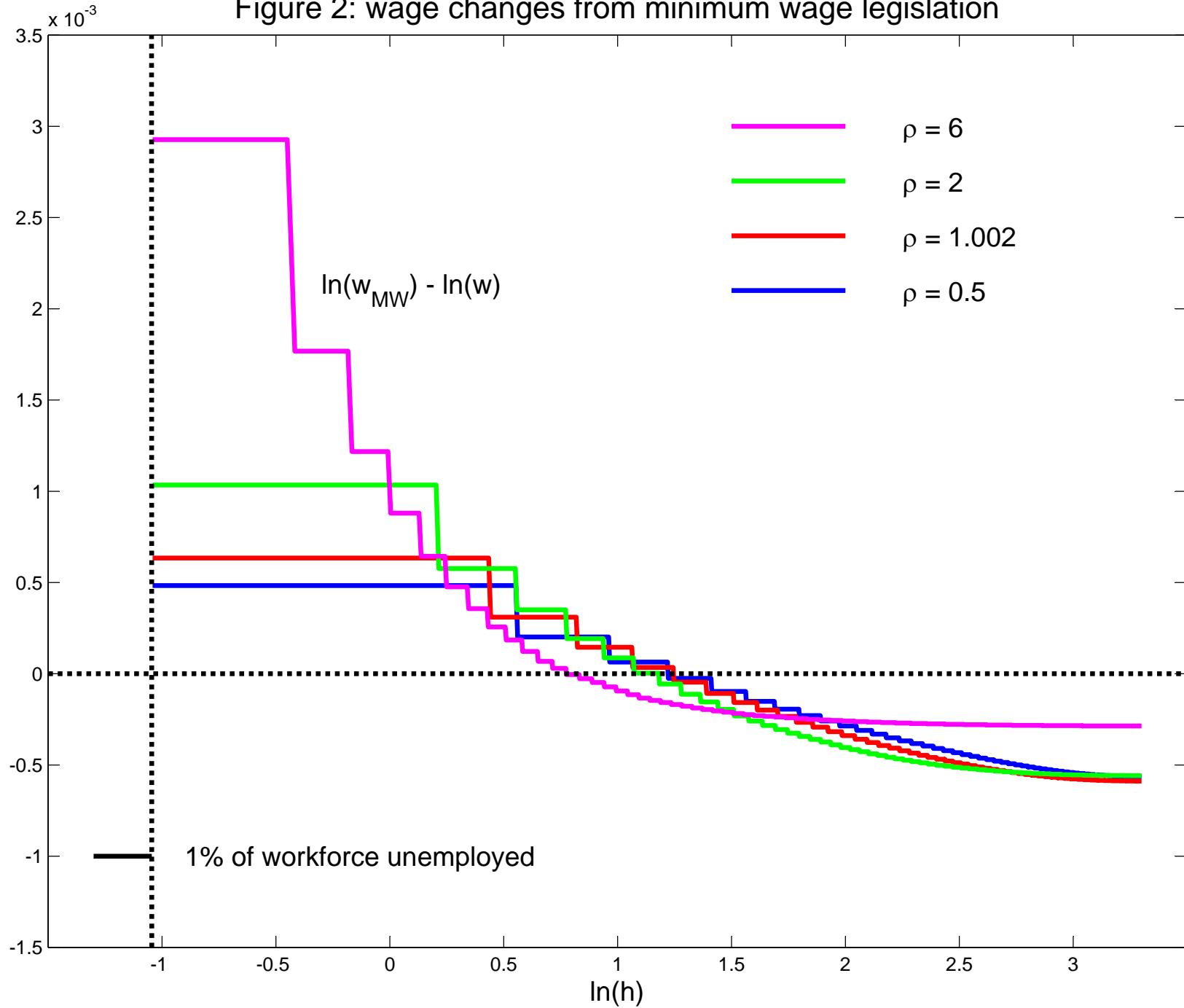
$\rho = 2$

$\rho = 1.002$

$\rho = 0.5$

1% of workforce unemployed

$\ln(h)$

Figure 3: wage changes from homogeneous technologies