

This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: The Impact of International Trade on Wages

Volume Author/Editor: Robert C. Feenstra, editor

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-23936-2

Volume URL: <http://www.nber.org/books/feen00-1>

Conference Date: February 27-28, 1998

Publication Date: January 2000

Chapter Title: And Now for Something Completely Different: An Alternative Model of Trade, Education, and Inequality

Chapter Author: Paul Krugman

Chapter URL: <http://www.nber.org/chapters/c6189>

Chapter pages in book: (p. 15 - 36)

And Now for Something Completely Different

An Alternative Model of Trade, Education, and Inequality

Paul Krugman

The dramatic growth of U.S. wage inequality since the early 1970s has been the subject of intense controversy both among economists and between economists and the broader public. To many people—including a few economists (e.g., Leamer 1998)—the explanation of that increased dispersion seems obvious: It is the result of globalization, and specifically of the growing imports of labor-intensive manufactures from developing countries. However, there are serious difficulties with an explanation of growing inequality that places the main weight on trade. For one thing, despite recent growth the value of north-south trade is still fairly small compared with the GDP of advanced economies; this means that even a complete elimination of that trade would, given reasonable estimates of factor shares and elasticities of substitution, reverse only a fraction of the observed change in wage differentials (Krugman 1995). Moreover, while trade can raise the relative demand for skilled labor by shifting the production mix toward skill-intensive sectors, in reality most of the rise in the relative employment of highly educated workers has taken place not via a change in the sectoral mix, but via a shift toward such workers within sectors—a shift that has taken place despite a sharp rise in education premia, which should have induced firms to substitute workers who were not college educated (Lawrence and Slaughter 1993).

But if globalization didn't do it, what did? An explanation that is consistent with the data is skill-biased technological change, taking place simultaneously in many sectors (and also presumably in many countries). And

Paul Krugman is the Ford International Professor of Economics at the Massachusetts Institute of Technology and a research associate of the National Bureau of Economic Research.

this is in fact the explanation that has been advanced by a number of economists, myself included, as the best available answer.

Yet even among those economists who believe that skill-biased technological change is the best explanation we have for the growing wage gap—certainly a better explanation than globalization—there is widespread uneasiness. This uneasiness stems both from the indirect nature of the evidence for such change—it is essentially inferred from the fact that the relative wages and the relative employment of the highly educated have moved in the same direction—and from the sense that technology is too much of a *deus ex machina*, something invoked to tie up the loose ends in our story rather than something we believe in on its own merits. I know that I am not alone in wondering whether all of us—both those who insist that globalization must be the explanation and those of us who regard the evidence against a simple Stolper-Samuelson account as overwhelming—are missing something, whether we may not all be on the wrong track.

The purpose of this paper is to offer a suggestive example of how a process quite distinct from either the simple trade or the simple technology story might be central to understanding growing inequality. The basic idea is that the labor market might, over some range of conditions, be characterized by multiple, locally stable equilibria, some more egalitarian than others. If that is the case, unequalizing shocks of modest size—shocks that could originate either in changing trade opportunities or in changing technology, or for that matter in both—could push the economy out of an egalitarian equilibrium and thus set in motion a cumulative process of growing inequality. In the specific model presented here, that process, which essentially feeds on itself, could easily be misinterpreted as exogenous skill-biased technical change.

The particular mechanism generating multiple equilibria in this model is a version of the screening/signaling hypothesis (Spence 1971; Stiglitz 1975). This approach was taken because it is the simplest labor market model with the required characteristics. It may, however, be only one of a number of possible mechanisms. For example, recent work by Acemoglu (1996), which is in somewhat the same spirit, offers a quite different mechanism involving technology choice in a frictional search model of the labor market. As we will see, there are some empirical difficulties with the specific model presented here; thus it should be considered only as a first exploration of a class of “exotic” income distribution models that might turn out to recast the nature of the debate.

The remainder of this paper is in five parts. Section 1.1 lays out the assumptions of the model. Section 1.2 develops the crucial idea of a distinction between two labor market regimes. Section 1.3 then shows how small shocks can precipitate a shift from one regime to the other, and thereby produce a cumulative process of growing inequality; it also shows how such a process might confuse someone trying to interpret the data.

Section 1.4 discusses some empirical predictions of the model, contrasting them with the predictions of alternative approaches and testing them loosely against the data. Finally, section 1.5 suggests some qualifications, extensions, and implications of the type of analysis this model represents.

1.1 Assumptions of the Model

We consider an economy endowed with two kinds of labor, “good” (G) and “bad” (B). The difference between these two kinds of workers is assumed to be inherent and unalterable. A worker’s type is known to the worker himself but is unobservable to employers.

Good workers do, however, have a way to demonstrate their goodness: They can acquire a college degree. It is assumed that only good workers are capable of acquiring such a degree, so the possession of a degree proves that a worker is of type G . However, acquiring a degree is costly. Rather than explicitly model this cost, I will simply assume that G workers will choose to become educated if and only if the ratio of the wage of those with degrees to those without, w_H/w_L , exceeds some value $r > 1$.

There are also two kinds of jobs: managerial (M) and nonmanagerial (N). M jobs actually require a college degree—that is, such a job can only be filled by a good worker with an education. In N jobs, there is no advantage in being educated per se. However, good workers are known to be $\pi > 1$ times as productive in N jobs as bad workers, so that even in N employment a worker with a college degree—who is therefore certified as type G —will command a higher wage than one without a degree.

Aside from this asymmetric-information feature of the labor market, the economy may be described by the standard two-by-two model of trade theory. There are two sectors, manager-intensive X and nonmanager-intensive Y . All individuals are assumed to share the same homothetic preferences over the two goods:

$$(1) \quad U = U(C_X, C_Y).$$

Production in each is described by a constant-returns function of managers and nonmanagers with all the usual properties:

$$(2) \quad Q_X = X(M_X, N_X),$$

$$(3) \quad Q_Y = Y(M_Y, N_Y).$$

In these production functions, M is simply the number of college-educated good workers employed as managers. N , however, must be measured in efficiency units because each good worker employed in a nonmanagerial role contributes π times as much as each bad worker. Thus the resource constraints for the economy are

$$(4) \quad M = M_X + M_Y = G_M,$$

$$(5) \quad N = N_X + N_Y = \pi G_N + B,$$

where G_M is the number of good workers employed as managers, and G_N is the number employed in nonmanagerial positions.

Finally, this is an open economy. It is not, however, a small open economy facing given world prices. Instead, it faces a less-than-perfectly-elastic rest-of-world offer curve. There are two reasons for using this large-open-economy setup. One is that it is arguably considerably more realistic than the price-taking assumption—for the United States alone, and certainly if we think of ourselves as modeling the Organization for Economic Cooperation and Development as a whole. More to the point for this paper, however, the large-economy setup is, for reasons that will soon become apparent, more convenient as a modeling device.

We will represent the rest-of-world offer curve by assuming that the relative price of X is decreasing in net exports of X :

$$(6) \quad \frac{P_X}{P_Y} = D(Q_X - C_X).$$

This completes the statement of the model. Next we turn to analysis.

1.2 Labor Market Regimes

In order to analyze this model, we need to determine the relationship between the number of good workers who acquire a college education—which we will denote by H —and the payoff of such an education, which we measure by the ratio of college to noncollege wages, w_H/w_L .

The nature of that relationship depends on the regime in the labor market, which can take one of two forms. First, it may be the case that college-educated workers are employed only in managerial positions—that is, in jobs for which such an education is actually necessary, in which education is actually socially productive. I will refer to this as the *human capital regime*. Alternatively, it may be the case that some college-educated workers are employed in nonmanagerial jobs. By assumption, the education does not enhance their productivity in these jobs; its only function is to demonstrate to employers that they are good as opposed to bad workers. I will refer to this as the *quality signaling regime*—although we must bear in mind that even under this regime there will be workers whose education actually is socially productive.

Before we determine the conditions under which each regime prevails, let us examine the behavior of the economy under each. Under the human capital regime, the number of college-educated workers is the same as the

number of managers, and the remaining workers form the supply of N in efficiency units:

$$(7) \quad M = H,$$

$$(8) \quad N = \pi(G - H) + B.$$

In terms of production and prices, then, the human capital regime constitutes an ordinary two-by-two economy, in which increases in the number of college-educated workers amount to an increase in M and a decline in N .

The properties of such an economy are very familiar. In particular, imagine either increasing M or decreasing N at constant relative goods prices. In either case, the economy would experience a Rybczynski effect: The output of M -intensive X would rise and the output of N -intensive Y fall. This would lead, at unchanged relative prices, to an increase in the excess supply of X . But given the rest-of-world offer curve (6), this means a decline in the relative price of X . So we may think in terms of a reduced-form relationship between the number of college-educated workers and relative prices:

$$(9) \quad \frac{P_X}{P_Y} = F(H).$$

But in such a two-by-two economy there is also a direct Stolper-Samuelson relationship between relative goods prices and relative factor prices. Thus we can write a reduced-form relationship between H and relative wages of the form

$$(10) \quad \frac{w_M}{w_N} = \phi(H), \quad \phi' < 0.$$

We are not, however, quite there yet. Equation (10) gives the ratio of the wage per manager (which is also the wage of any college graduate) to the wage per efficiency unit of nonmanagerial labor. However, because some of the workers in nonmanagerial jobs are (unidentified) type G , the average number of efficiency units per N worker is more than one; so the wage paid per actual worker in N jobs is

$$(11) \quad w_L = w_N \frac{\pi(G - H) + B}{G - H + B}.$$

Notice that the ratio of w_L to w_N is decreasing in H . The reason is that the more G workers who acquire college educations, the lower the expected productivity of the average worker without a degree. This complicates matters in the human capital regime, but is central to the story in the quality signaling regime.

The education premium as a function of H is therefore

$$(12) \quad \frac{w_H}{w_L} = \phi(H) \frac{G - H + B}{\pi(G - H) + B}.$$

This relationship can, in principle, be either downward- or upward-sloping. It will be downward-sloping if the effect of factor supplies on relative prices is strong (this effect would be nonexistent, of course, in a price-taking economy—which is why the large-economy assumption is useful) and if the screening effect is weak. I will assume that this is in fact the case, so that the curve relating w_H/w_L to H in the human capital regime looks like curve HC in figure 1.1.

It may be useful at this stage to give a specific example. Suppose that we consider a closed economy—the limiting case of a large open economy—in which both tastes and technology are Cobb-Douglas. Let μ be the share of X in expenditure, and let α and β be the share of M in the X and Y sectors, respectively. It follows immediately that a share $\mu\alpha + (1 - \mu)\beta$ of total income accrues to M and a share $\mu(1 - \alpha) + (1 - \mu)(1 - \beta)$ to N , so that

$$(13) \quad \frac{w_M M}{w_N N} = \frac{\mu\alpha + (1 - \mu)\beta}{\mu(1 - \alpha) + (1 - \mu)(1 - \beta)},$$

and thus

$$(14) \quad \frac{w_M}{w_N} = \frac{\mu\alpha + (1 - \mu)\beta}{\mu(1 - \alpha) + (1 - \mu)(1 - \beta)} \frac{N}{M} = \kappa \frac{\pi(G - H) + B}{H}.$$

Now substituting in equation (11), we find that

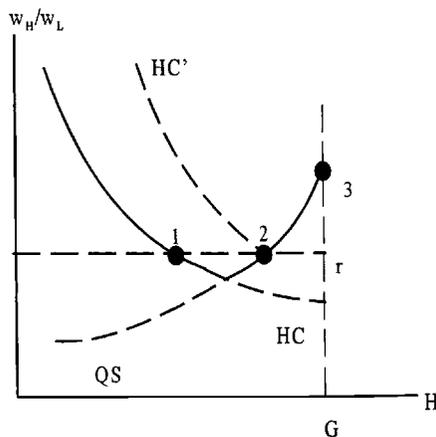


Fig. 1.1 Determining skills and wages

$$(15) \quad \frac{w_H}{w_L} = \kappa \frac{G - H + B}{H},$$

which is unambiguously a downward-sloping curve.

Next let us turn to the quality signaling regime. This is more straightforward. Since (some) college-educated workers are competing directly with non-college-educated workers, each college-educated worker will receive a wage equal to π times the wage per efficiency unit in N , while each noncollege worker will receive a wage proportional to the average number of efficiency units among such workers. But the total number of efficiency units among noncollege workers is

$$(16) \quad E = \pi(G - H) + B,$$

so the average efficiency is

$$(17) \quad \frac{E}{L} = \frac{\pi(G - H) + B}{G - H + B},$$

implying the relative wage equation,

$$(18) \quad \frac{w_H}{w_L} = \pi \frac{G - H + B}{\pi(G - H) + B},$$

which is unambiguously upward-sloping, yielding a curve like QS in figure 1.1.

Which regime prevails? The answer, of course, is whichever provides the higher relative wage. If the HC curve lies above the QS curve for some given H , college graduates can command higher wages in managerial than in nonmanagerial jobs, even if all graduates are so employed; so no college-educated workers will use their education purely to demonstrate their type. If HC lies below QS , then if all graduates worked as managers they would earn less than a certified good worker could get in a nonmanagerial job, and so some H workers will shift away from the managerial role. The overall relationship between H and w_H/w_L , then, is V-shaped; it is defined by the upper envelope of the curves HC and QS .

1.3 Equilibria and Transitions

Let us now endogenize the supply of college-educated workers. We have assumed that good workers will find it worthwhile to acquire a college education if and only if w_H/w_L exceeds some value $r > 1$. Thus in figure 1.1, which represents the most interesting case, the solid line (the upper envelope of the HC and QS curves) represents the demand for college graduates, while the broken horizontal line at $w_H/w_L = r$ represents the supply.

Clearly, there are three equilibria: one at point 1, where only some good

workers get degrees and the labor market is in a human capital regime; one at point 3, where all those who can get degrees do and the quality signaling regime prevails; and an intermediate equilibrium at point 2. Which equilibrium does the economy select? For some purposes it would be essential to model the full-fledged dynamics of educational investment, including the formation of expectations about future returns. However, in recent years the interest in evolutionary game theory suddenly has made ad hoc dynamics, in which the mix of strategies followed by a population gradually changes depending on the current returns to each strategy, respectable again. Such evolutionary dynamics are particularly useful when, as in this model, there are multiple equilibria and we are looking for a selection criterion. Or to make a long story short, it is acceptable as a first cut to assume that the number of college-educated good workers rises if $w_H/w_L > r$, and falls if $w_H/w_L < r$. The equilibrium at point 2 is, then, unstable, while the other equilibria are locally stable. Alternatively, if one thinks of the strategy of good workers as probabilistic (e.g., with what probability will I go to college?), then points 1 and 3 are evolutionarily stable strategies in the sense of Maynard Smith (1982), while point 2 is not.

Where the economy ends up, then, depends on history. In particular, if it manages to get into the relatively egalitarian equilibrium at point 1, it will tend to stay there in the face of small shocks.

But now suppose that there is a progressive rise in the demand for managerial workers. This rise could be due either to growing opportunities to export the manager-intensive good X , or to manager-biased technical change. In either case, the effect under the human capital regime will be to increase the relative wage of H workers for any given level of H , an upward shift in the HC curve.

As long as this shift is not too large, the labor market will remain in the human capital regime, and the wage differential will not change. At a critical point, however, illustrated in figure 1.1 by the curve HC' , the human capital equilibrium will cease to exist. At this point, even if there is no further increase in the demand for managerial workers, the economy will continue evolving toward increasing inequality. As a growing fraction of good workers become educated, the expected productivity and thus the wages of non-college-educated workers will fall, further increasing the incentive to acquire a degree, continuing until the unequal equilibrium at point 3 has been reached.

Two observations need to be made about this transition to higher inequality. The first is that if anything like this story is right, the whole attempt to apportion the causes of growing wage differentials between technology and trade may be missing the point. Either trade or technology—or more likely both—may push the economy to the critical point, but thereafter the unequalizing process simply feeds on itself, and the proximate cause may therefore be irrelevant.

The second observation is how easily such a transition might be misin-

terpreted by an observer—me, for example—who works with the wrong model. Suppose that exogenous forces in fact were to push HC just to the critical point and that the economy were then to evolve spontaneously from point 2 to point 3. What would we see? Because of the strong structure of this model, we can immediately determine that there will be no change at all in the wages of managers relative to the cost of one efficiency unit of nonmanagerial work. We know this because all college graduates under the quality signaling regime receive a wage equivalent to π units of N . But if these true relative factor prices are unchanged, relative goods prices must also remain unchanged, and so therefore must production, consumption, and trade. *Nothing real changes as the economy moves from point 2 to point 3, except the distribution of wages.*

But an observer who classifies labor not by unobservable quality but by education level will see two things happening: a fall in both the relative and the absolute wages of non-college-educated workers, and a fall in their relative employment. This seemingly perverse outcome will not appear to be explained by trade—neither the volume of trade nor the prices at which trade takes place will change, nor will the industry mix of employment; all of the increased demand for H and reduced demand for non- H will therefore come from within-sector shifts. The only hypothesis that will appear to be consistent with the data will therefore be skill-biased technical change—even though no technical change has in fact taken place!

1.4 Empirical Implications

At the beginning of this paper, I described some reasons why even those of us who have provisionally adopted a technology story for growing wage inequality nonetheless feel uneasy. But beyond gut feelings, is there some way to distinguish among the alternative explanations of inequality?

Suppose for a moment that we try to explain rising wage differentials with one of three pure stories: Stolper-Samuelson effects with no change in technology, skill-biased technological change with no change in factor supplies, and a pure story about transition from a human capital to a quality signaling regime (that is, the movement from point 2 to point 3 in figure 1.1). In reality, of course, there is no reason why the data should be generated by a pure story; but this comparison is nonetheless illuminating. Let us, therefore, compare the implications for three observable variables.

Skill Intensity within Industries. The original exposition of Stolper and Samuelson relied on the device of a contract curve within an Edgeworth box to make the now familiar, although sometimes still misunderstood, point that trade affects the demand for factors via its effect on the industry mix; in their analysis, a rise in the relative price of the labor-intensive good causes employment to shift toward that good. Full employment of factors is preserved via a compensating shift toward capital-intensive techniques

within each sector. Thus in a pure Stolper-Samuelson account of rising wage differentials, we would expect to see the skill intensity of production falling within each industry.

Skill-biased technological change, in contrast, would tend to raise the relative demand for skilled labor within each industry. If we take factor supplies as given, however, the skill premium would have to rise enough to choke off this shift in relative demand in the aggregate; while the precise effects on each industry would depend on both the distribution of technological change and the elasticity of substitution, on average there would be no change in skill intensity within industries. (Of course, if we imagine that relative supplies of skilled labor are increasing at the same time, changes in factor prices would be less and we would therefore see increases in skill intensity within industries.)

Finally, in the transition from a human capital to a quality signaling regime, nothing real would change, but firms in both sectors would begin to employ college-educated workers for previously unskilled jobs; thus the observed skill intensity would rise in all sectors.

Factor Prices. It is a fundamental proposition in trade theory that following a change in relative prices, the bundle of factors initially employed must be at least able to afford the goods they were previously producing. (This proposition is the basis for the demonstration of gains from trade in terms of the dual; see Helpman and Krugman 1985, chap. 1.) Thus, a Stolper-Samuelson explanation of changing factor prices implies that the purchasing power of the original employment of factors in terms of output (or, given the absence of strong terms of trade effects, in terms of consumption) must rise or at least not fall.

The same is a fortiori true of an explanation in terms of technological change: Technological progress must increase the purchasing power of the initial bundle of factors for any given goods prices, and any change in goods prices can only further increase that purchasing power.

A pure transition from a human capital to a quality signaling regime, however, actually reduces the purchasing power of the initial bundle of factors, if workers are classified by education. In the transition from point 2 to point 3 in figure 1.1, neither the average wage rate nor the wage rate of skilled workers changes; but the wage rate of workers without a college education falls. Thus an index with fixed weights on the college-educated and non-college-educated wage will unambiguously decline.

Total Factor Productivity. Roughly speaking, this is dual to the measurement of factor purchasing power. If total factor productivity (TFP) is properly measured, it should not change at all in a pure Stolper-Samuelson story. It should, of course, rise if there is technological progress, so it must increase in a technology-driven account. Again, a transition from a human capital to a quality signaling regime should have a perverse implica-

tion if labor types are measured by education: There will be an increase in human capital but no increase in output, so TFP will (as measured) actually fall.

We may therefore summarize the predictions of the two standard approaches to inequality and of the exotic alternative offered here in table 1.1.

What do we observe in practice? It is a familiar point since the work of Lawrence and Slaughter (1993) that in the United States there has been a pervasive shift toward higher skill intensity within industries, with relatively little shift of the industry mix of employment toward skill-intensive products. Thus on this first criterion, the two conventional approaches seem to fail (although the technology-driven story can be rescued by supposing that factor supplies have exogenously shifted, albeit not as rapidly as factor demands).

When we come to the purchasing power of factors, what we observe for the United States over the period of rising inequality is a slow rise in average real wages (returns on capital are more problematic to measure). However, this rise in real wages has been accompanied by a rising average educational level, so that it is unclear whether the real income of the initial basket of factors has increased. Table 1.2 presents a calculation using data from Mishel (1996). The first column shows the 1979 share of workers by educational category; since the total does not add up to 100, the second column prorates the difference to derive an adjusted share. The third column shows the real hourly wage rate (in 1996 dollars) of each worker type in 1979; the fourth the real wage in 1989. As the last line of the table indicates, according to these numbers the average real wage weighted by

Table 1.1 Predictions of Alternative Approaches Based on Change

	Skill Intensity within Industries	Purchasing Power of Initial Factors	Total Factor Productivity
Stolper-Samuelson	Negative	Non-negative	Zero
Skill-biased technology	None	Positive	Positive
Transition to screening	Positive	Negative	Negative

Table 1.2 Real Wages of a Constant-Skill Basket

	1979 Share of Workforce	Adjusted Share	Real Wage, 1979	Real Wage, 1989
Non-high school	28.5	29.2	10.59	8.91
High school	41.7	42.7	11.86	10.79
Some college	15.1	15.5	12.92	12.53
College	8.8	9.0	16.55	16.98
Advanced degree	3.6	3.7	20.34	22.07
Weighted average			12.39	11.48

Table 1.3 Factors Accounting for U.S. Growth

	Output per Worker	Capital per Worker	Human Capital per Worker	Total Factor Productivity
1960-94	1.1	0.4	0.4	0.3
1960-73	1.9	0.5	0.6	0.8
1973-94	0.6	0.3	0.2	0.1
1973-84	0.2	0.3	0.5	-0.5
1984-94	0.9	0.3	0.0	0.7

1979 labor force composition actually declined approximately 7 percent during the decade of the 1980s. While this rough calculation might not stand up in the face of a careful cleaning up of the data, it is remarkable that such a fixed-weight wage index falls despite technological progress, a roughly constant labor share in GDP, and stable or improving terms of trade. The result is, at least on the face of it, inconsistent either with Stolper-Samuelson or technology-driven stories; it is consistent with the transition to a screening equilibrium.

As one might expect given its rough conceptual equivalence, calculations of total factor productivity yield similarly puzzling results. Table 1.3 shows a typical recent calculation from Collins and Bosworth (1996); it suggests that during the post-1973 period, despite what looks like continuing technological advance, growth in total factor productivity ground to a near halt. Again, a possible, although not necessarily correct, explanation is that the accumulation of human capital over that period represented a socially unproductive investment in screening.

To me, at least, this rough evidence clearly indicates not only the well-established point that Stolper-Samuelson effects cannot have been the main driving force behind changes in the wage distribution, but also that something funny is going on that is not easily mapped into a simple technology-driven story either. In particular, the aggregative implications of the transition to screening seem to fit the actual data quite well.

However, we should note that there is one important implication of the particular type of model developed here that appears to be contradicted by the data. If the multiple equilibria arise only because workers are better sorted by quality than they used to be, one ought to observe a difference in the growth of inequality by cohort: The skill premium for workers from older cohorts, who made educational decisions at a time when screening motives were not as important, should not have increased at the same rate as that for more recent cohorts (i.e., a worker who chose not to go to college in 1960 revealed less about himself than a worker making the same choice in 1975). In fact, however, cohort studies (e.g., Juhn, Murphy, and Pierce 1993) suggest comparable increases in inequality across cohorts.

1.5 Qualifications, Extensions, and Implications

It should go without saying that the model proposed here is very special, and that there are good reasons to be skeptical about the mechanism proposed. In general, signaling/screening models of the labor market have been questioned by many labor economists who wonder why employers would not attempt to create cheaper sorting mechanisms and spare good workers the huge expense of pointless college attendance. More generally, this model suggests that what looks like skill-biased technological change might actually be something else; but then again it might be skill-biased change after all.

On the other hand, the mechanism described here is only one of a class of possible positive-feedback stories about growing inequality. As mentioned in section 1.1, Acemoglu (1996) offers an alternative story roughly along the following lines. In a search economy, in which neither firms nor workers can expect to find an immediate match, firms must choose whether to implement a technology that is highly productive only if a skilled worker uses it, or a more robust technology that any worker can use. Firms will have an incentive to implement the skill-sensitive technology if they can quickly find skilled workers, so the demand curve for such workers will be upward- rather than downward-sloping. (This story does not suffer from the objection that it should apply only to younger cohorts of workers.) Other economists have suggested that since the power of an individual union depends in part on the strength of a national union movement, the dramatic decline in union membership in the United States in recent decades may represent a self-reinforcing process contributing to inequality. One might even invoke linkages that run through the political economy of policy; for example, Benabou (1996) has proposed that, over some range, increased inequality tends to lead to policies that reinforce that inequality. And no doubt there are other possible mechanisms to be considered—nor need such stories be mutually exclusive.

If anything like the mechanism proposed in this paper is indeed at work, the policy implications cut sharply across all current orthodoxies. Consider, for example, the implications for trade policy. It is possible in this kind of model that globalization could be the proximate cause of a process that then gives the false appearance of being driven by exogenous skill-biased technical change. This does not mean, however, that the process could be reversed by reversing the globalization: Even if trade tipped the balance and undermined the human capital equilibrium, once the economy is in a quality signaling equilibrium, a small downward shift in the *HC* curve will not push it out again. So one could blame trade for increased inequality, yet at the same time conclude that protectionism cannot do much to reduce wage differentials.

Or consider the popular proposals of Reich (1991) and others to combat inequality by promoting the acquisition of human capital—in effect, by subsidizing education. In this model, such policies would be completely ineffective once the economy is in the quality signaling regime. Indeed, they could actually be counterproductive if an economy is still in a human capital regime, but close to the critical point: Anything that encourages good workers to get educated can set in motion a cumulative process of growing inequality!

In short, while the specific model presented in this paper is implausible in its details, it may be very important as a practical matter to contemplate the possibility that the real causes of growing inequality are very different from any of the explanations that have dominated recent debate.

References

- Acemoglu, D. 1996. Changes in unemployment and wage inequality: An alternative theory and some evidence. Cambridge, Mass.: Massachusetts Institute of Technology. Mimeograph.
- Benabou, R. 1996. Inequality and growth. In *NBER macroeconomics annual 1996*, ed. B. Bernanke and J. Rotemberg. Cambridge, Mass.: MIT Press.
- Collins, S., and B. Bosworth. 1996. Economic growth in East Asia. *Brookings Papers on Economic Activity*, no. 2: 135–91.
- Helpman, E., and P. Krugman. 1985. *Market structure and foreign trade*. Cambridge, Mass.: MIT Press.
- Juhn, C., K. M. Murphy, and B. Pierce. 1993. Wage inequality and the rise in returns to skills. *Journal of Political Economy* 101:410–42.
- Krugman, P. 1995. Growing world trade: Causes and consequences. *Brookings Papers on Economic Activity*, no. 1: 327–62.
- Lawrence, R., and M. Slaughter. 1993. International trade and U.S. wages in the 1980s: Giant sucking sound or small hiccup? *Brookings Papers on Economic Activity, Microeconomics*, no. 2: 161–210.
- Lerner, E. 1998. In search of Stolper-Samuelson linkages between international trade and lower wages. In *Imports, exports, and the American worker*, ed. Susan Collins, 141–214. Washington, D.C.: Brookings Institution.
- Maynard Smith, J. 1982. *Evolution and the theory of games*. Cambridge: Cambridge University Press.
- Mishel, L. 1996. *The state of working America*. Washington, D.C.: Economic Policy Institute.
- Reich, R. 1991. *The work of nations*. New York: Basic Books.
- Spence, A. M. 1971. *Market signalling*. Cambridge, Mass.: Harvard University Press.
- Stiglitz, J. 1975. The theory of “screening,” education, and the distribution of income. *American Economic Review* 65:282–300.

Comment James E. Rauch and Magnus Lofstrom

We begin with the empirical part of our discussion. We take the basic empirical message of Krugman's paper to be that the rise in the education wage premium is due to the choice by an increasing number of good workers to get a college education, which allows them to be distinguished from "bad" workers and therefore earn a higher wage. This lowers the average quality of noncollege-educated workers and, thus, raises the difference between the wage of the educated and the noneducated. We examine the implications of Krugman's hypothesis for the contrast between wage/salary earners and the self-employed. The latter presumably earn their true productivity rather than being pooled by their employers.

If education is only a signal and markets are perfect, there will be no education-earnings differential for self-employed of equal ability. This might be true for individuals running small retail/wholesale trade or manufacturing establishments. On the other hand, the self-employed who provide professional services benefit from having credentials. A self-employed accountant or stockbroker may only need innate ability and high school math, but has to get some formal education to obtain credentials (e.g., a CPA degree). The earnings of these individuals are likely to be set by competition from large accountant and stockbroker firms. Suppose these individuals have exactly 16 years of schooling (we discuss postgraduate education later). In this case, when we look at the earnings differential between self-employed workers with 16 years of schooling and 12 years of schooling, we see a mix of a group for which the good agents have no incentive to sort themselves by obtaining an education¹ and a group where good agents are able to earn a credential (equivalent to a signal); hence, we should see a smaller wage differential than for wage/salary earners, according to Krugman's model. Over time, as more of the good individuals choose to get a credential (signal), the education-wage differential increases for the self-employed as a whole, but not as much as for wage/salary earners. Moreover, the earnings of the self-employed with 12 years of education increase relative to wage/salary earners with 12 years of education because the overall quality of the former group declines less.²

We will perform crude tests of these hypotheses by examining log weekly earnings for wage/salary earners and the self-employed in 1980 and 1990. Our sample consists of males ages 18 to 64 and is drawn from the

James E. Rauch is professor of economics at the University of California, San Diego, and a research associate of the National Bureau of Economic Research. Magnus Lofstrom is a research associate of the Institute for the Study of Labor (IZA) in Bonn, Germany.

1. Some may have obtained an education anyway, for its consumption value or perhaps because they anticipated wage/salary employment and switched to self-employment later in their working lives.

2. We are indebted to Alan Krueger for this last observation.

Table 1C.1 College–High School Earnings Differentials, Wage/Salary Earners versus Self-Employed

	Difference, 16 and 12 Years Education Mean Log Wage
1. Wage/salary	
1980	0.3211227
1990	0.4696073
Self-employed	
1980	0.3590479
1990	0.4299948
2. Wage/salary	
1980	0.3183041
1990	0.4653192
Self-employed	
1980	0.2719227
1990	0.3574143

Note: For row 1, earnings = wage/salary income plus self-employment income for both groups. For row 2, earnings = wage/salary income for wage/salary workers and self-employment income for the self-employed.

U.S. census. Additional details are given in the appendix. Individuals are self-identified as wage/salary earners or self-employed. In tables 1C.1–1C.4, we will first use total earnings (wage/salary earnings plus self-employment earnings) for both groups and then use wage/salary earnings only for wage/salary earners and self-employment earnings only for the self-employed. We feel the latter earnings measures are more appropriate for testing the hypotheses.

In table 1C.1 we see that the earnings differential between males with 16 versus 12 years of education is indeed higher for wage/salary earners than for the self-employed in all cases except for total earnings in 1980. More importantly for Krugman's main point, the education earnings differential increased more between 1980 and 1990 for wage/salary earners than for the self-employed regardless of which earnings measure is used. In table 1C.2 we see that the earnings of the self-employed with 12 years of education improved compared to their wage/salary-earner counterparts between 1980 and 1990. All of these observations are consistent with Krugman's signaling model of increased earnings inequality.

Often the rising wage inequality in the United States is cast as a rising return to ability. Thus, within demographic groups inequality has risen steadily since the late 1960s, as Katz and Murphy (1992) show. Table 1C.3 shows no such rise among the self-employed during the 1980s when we control for a quartic in experience and years of education interacting with four education-level dummies, as in Katz and Murphy (1992, 44). As we

Table 1C.2 Earnings Differentials between Wage/Salary and Self-Employed Workers with 12 Years of Education

Earnings	Difference in Mean Log Wage between Self-Employed and Wage/Salary Earners
Wage/salary income + self- employment income	
1980	0.0647802
1990	0.1164628
Wage/salary income for wage/salary workers and self-employment income for the self-employed	
1980	-0.1210538
1990	-0.0534281

Table 1C.3 Earnings Inequality for Wage/Salary Earners versus Self-Employed

	Difference, 90th and 10th Percentiles		Variance	
	Actual	Residual	Actual	Residual
1. Wage/salary				
1980	1.346599	1.17467	0.329445	0.25176
1990	1.579453	1.30136	0.394088	0.28254
Self-employed				
1980	2.239694	1.94543	0.693338	0.59222
1990	2.415007	2.02122	0.725025	0.59407
2. Wage/salary				
1980	1.35460	1.17005	0.3263	0.24984
1990	1.57168	1.29971	0.3892	0.28031
Self-employed				
1980	2.25039	2.02838	0.6832	0.60456
1990	2.10476	1.90466	0.6399	0.54119

Note: For row 1, earnings = wage/salary income plus self-employment income for both groups. For row 2, earnings = wage/salary income for wage/salary workers and self-employment income for the self-employed.

expect, the level of inequality among the self-employed is much greater than for wage/salary earners, reflecting greater scope for ability in self-employment. Our interpretation of table 1C.3 is that it supports Krugman's claim that nothing "real" happened to change the return to ability during the 1980s and that, instead, the increase in ability differentials among wage/salary earners probably reflects the breakdown of internal labor market rigidities, as suggested by Reich (1991).

Finally, we return to the education-wage differential, but include males

Table 1C.4 College and Postgraduate–High School Earnings Differentials, Wage/Salary Earners versus Self-Employed

	Difference in Mean Log Wage between 16+ and 12 Years Education
1. Wage/salary	
1980	0.3819170
1990	0.5549449
Self-employed	
1980	0.5212091
1990	0.6658614
2. Wage/salary	
1980	0.3767991
1990	0.5477357
Self-employed	
1980	0.4705068
1990	0.5957573

Note: For row 1, earnings = wage/salary income plus self-employment income for both groups. For row 2, earnings = wage/salary income for wage/salary workers and self-employment income for the self-employed.

with postgraduate education. This group includes, for example, dentists and doctors who really could not perform their jobs without formal training. We see from table 1C.4 that the relationship between the earnings differential for wage/salary earners and the self-employed is essentially reversed compared to table 1C.1. This suggests that postgraduate education is truly productive and that there is greater scope for realizing returns to this productive education in self-employment. The increase in the education earnings premium for the self-employed is still smaller than for wage/salary earners, however.

We now turn to the theoretical part of our discussion. We find the weakest element of Krugman's model to be the human capital regime, in which the increased relative supply of educated workers drives down their relative wages. Krugman gets around the factor-price insensitivity theorem (Leamer 1995) by allowing relative prices to change as a result of the United States being a large country. But this implies that a glut of college-educated workers in the United States, such as occurred in the 1970s, should lower returns to education everywhere else in the world to the same extent as in the United States. The real problem with traditional models is not the insensitivity of relative factor prices to changes in relative supplies of different kinds of factors, but rather insensitivity of relative factor prices for the same factor across countries to changes in relative country supplies. In other words, the relative country labor-demand curve is infinitely elastic for any given type of labor—the relative country wage is fixed by price competition at the margin.

One can solve this problem with traditional models by using different cones of specialization or moving beyond the two-by-two case. In keeping with the spirit of Krugman's paper, however, let us consider a model with incomplete information, specifically that of Rauch and Casella (1998). In this model, information is incomplete in product markets rather than labor markets. Incomplete information is modeled as a matching problem: It is more difficult in international than in domestic markets for producers to find the right distributors for their consumer goods, assemblers to find the right suppliers for their components, investors to find the right partners for their joint ventures, and so on. Production takes place as follows: Producers match pairwise, and, if the match is acceptable, labor is employed to realize the productive opportunity. Domestic matching takes place subject to complete information—every producer knows the type of every other producer—while international matching is effectively random. However, each producer has access only to the labor in his own country, so domestic matches must employ domestic labor, while international matches can employ labor in whichever country it is cheapest. International matches can thus serve to transfer labor demand from the country where labor is scarce to where it is abundant—the price system at work.

Although an informational barrier, such as a conventional trade barrier, is uniform across producers *ex ante* in this model, some producers may effectively evade it *ex post* if they are lucky in their international matching. If enough producers match successfully abroad, sufficient labor demand is transferred to equalize wages in the two countries. If the two countries' producer-labor endowment ratios are far enough apart, however, lucky or good matches cannot transfer sufficient labor demand to equalize wages. Rather than being constant, as with a traditional trade barrier, the proportional wage gap will increase as the producer-labor endowment ratios move further apart: The relative labor-demand curve becomes downward-sloping. However, a country's labor force is not fully insulated from changes in the labor endowment of its trading partner. As we move down the relative labor-demand curve, some bad matches become acceptable because of the gains from trade provided by the increased wage differential, generating an additional transfer of labor demand. Thus an increase in one country's endowment of labor decreases wages in both countries, but decreases its own wage more.³

This model offers a resolution to the "margins versus volumes" debate in the trade and wages controversy: Are choices made at the margin sufficient to equalize factor prices across various national markets, or is a certain volume of trade across the markets necessary to drive prices to equal-

3. In traditional models, a trade barrier that acts like a quota permits the equalization of wages when the quota is not binding, but when it is binding changes in countries' endowments can only affect their own wages.

ity? In a perfectly competitive world with no barriers to trade of any kind, this question is not well posed; if prices did not equalize, the volume of trade required to equalize them would automatically occur. In the real world, the debate has been elegantly summarized by Freeman (1995, 21–22):⁴

If the West can import children's toys produced by low-paid Chinese workers at bargain basement prices, surely low-skilled westerners, who produce those toys at wages 10 times those of the Chinese, will face a difficult time in the job market. It isn't even necessary that the West import the toys. The threat to import them or to move plants to less-developed countries to produce the toys may suffice to force low-skilled westerners to take a cut in pay to maintain employment. In this situation, the open economy can cause lower pay for low-skilled workers even without trade: to save my job, I accept Chinese-level pay, and that prevents imports. The invisible hand would have done its job, with proper invisibility. . . . These predictions [of factor-price equalization] run counter to a wide body of evidence that domestic developments do affect wages: for instance, that the baby boom affected the pay of young workers; that the relative number of college graduates altered the premium paid for education. . . .

Having recognized the theoretical point that competition from labor in low-wage countries could affect wages of comparably skilled labor in high-wage countries even in the absence of any actual trade, the empirical method for quantifying such competition that is preferred by Freeman and many others (see, e.g., Sachs and Shatz 1994) remains factor content analysis: "if the United States imported 10 additional children's toys, which could be produced by five American workers, the effective supply of unskilled workers would increase by five. . . . This five-worker shift in the supply-demand balance would put pressure on unskilled wages to fall, causing those wages to fall in accord with the relevant elasticity" (Freeman 1995, 23). Here the impact of low-wage competition depends entirely on the volume of net trade and not at all on the comparison of prices at the margin.

In the model of Rauch and Casella under discussion, if the volume of international matches that are acceptable without wage differentials (the good matches) is sufficiently large, margins operate perfectly: Any incipient rise in the labor-scarce country's wage compared to the labor-abundant country's wage can be eliminated by a shift in the labor demanded by these internationally matched producers from the former to

4. We might take the liberty of clarifying Freeman's statement to state that the baby boom affected the pay of young workers *more in the United States than in China* and the relative number of college graduates *in the United States* altered the premium paid for education *more in the United States than in China*.

the latter country. But if the volume of these good international matches is insufficient, margins operate imperfectly: The elimination of an incipient increase in the labor-scarce country's wage effectively requires that this country's producers shift from matching in a complete-information environment domestically to an incomplete-information environment abroad. When margins operate imperfectly, relative wages become a downward-sloping function of relative labor supplies in the two countries. In short, the efficient operation of margins is dependent on the adequate volume of good international matches, where the volume of these matches that is adequate is determined by the ratio of the factor endowments of the two countries. The actual volume of good matches is in turn determined by some combination of what could loosely be described as the model's information or matching technology and familiarity between the trading partners.

Appendix

Sample Selection Criteria and Computation of Weekly Earnings

Data are drawn from the 1980 5 percent A sample and the 1990 5 percent sample of the U.S. Census of Population. All immigrants and a 20 percent random subsample of natives who satisfy the following criteria are included: males between the ages of 18 and 64; not residing in group quarters; not in military service; not enrolled in school; reported working at least 40 weeks; earned at least \$100 per week in 1989 dollars. Weights are used and adjusted accordingly (i.e., weights for natives are multiplied by 5). Earnings reported in tables 1C.1–1C.4 are weekly wages. These are calculated by dividing annual earnings by the number of weeks worked in the year prior to the census. The number of observations range from 4,498 for self-employed males with 16 years of education in 1980 to 538,029 for wage/salary workers in 1990.

References

- Freeman, Richard B. 1995. Are your wages set in Beijing? *Journal of Economic Perspectives* 9 (summer): 15–32.
- Katz, Lawrence F., and Kevin M. Murphy. 1992. Changes in relative wages, 1963–1987: Supply and demand factors. *Quarterly Journal of Economics* 107 (February): 35–78.
- Leamer, Edward E. 1995. The Heckscher-Ohlin model in theory and practice. Princeton Studies in International Finance no. 77. Princeton, N.J.: Princeton University, Department of Economics, International Finance Section, February.

- Rauch, James E., and Alessandra Casella. 1998. Overcoming informational barriers to international resource allocation: Prices and group ties. NBER Working Paper no. 6628. Cambridge, Mass.: National Bureau of Economic Research, June.
- Reich, Robert. 1991. *The work of nations*. New York: Basic Books.
- Sachs, Jeffrey D., and Howard J. Shatz. 1994. Trade and jobs in U.S. manufacturing. *Brookings Papers on Economic Activity*, no. 1: 1–84.