International Trade and the Domestic Wage Structure
Alternative Explanations

Vivek H. Dehejia
International Trade and the Domestic Wage Structure
Alternative Explanations

Vivek H. Dehejia

Reihe Ökonomie / Economics Series No. 34

September 1996

Vivek H. Dehejia
The Norman Paterson School of International Affairs
Carleton University
1125 Colonel By Drive
Ottawa, CANADA K1S 5B6
Phone: ++1 (613) 520-6661
Fax: ++1 (613) 520-2889
e-mail: vdehejia@ccs.carleton.ca

Institut für Höhere Studien (IHS), Wien
Institute for Advanced Studies, Vienna
The Institute for Advanced Studies in Vienna is an independent center of postgraduate training and research in the social sciences. The Economics Series presents research done at the Economics Department of the Institute for Advanced Studies. Department members, guests, visitors, and other researchers are invited to contribute and to submit manuscripts to the editors. All papers are subjected to an internal refereeing process.

Editorial
Main Editor:
Robert M. Kunst (Econometrics)
Associate Editors:
Christian Helmenstein (Macroeconomics)
Arno Riedl (Microeconomics)
Abstract

The paper considers the "trade and wages" debate, and proposes two alternative explanations to explain the rising wage differential (relative wage of the skilled vs. the unskilled), other than the conventional Stolper-Samuelson explanation. The first is an explanation dubbed "kaleidoscopic comparative advantage": the argument is that increased labour turnover might differentially impede the human capital accumulation of the unskilled as against the skilled, leading to an alternative trade-based explanation. The second explanation is "capital-skill complementarity", drawing on the well-established empirical regularity that capital and skill are complementary with each other vis-à-vis unskilled labour. The paper builds a dynamic model, embedding this insight into a neoclassical adjustment model, and traces out the transitional dynamics and steady state.

Keywords
Wage inequality trade, skill-biased technical change

JEL-Classifications
F10, J31
Comments
This paper is based on Chapter 2 of my Columbia dissertation. I am grateful to Jagdish Bhagwati, my dissertation sponsor, for his unstintingly generous guidance and support, to R.L. Carson and Alan Deardorff for valuable comments on an earlier draft, and to Douglas W. Dwyer for his encouragement and invaluable technical assistance. The paper was presented at the Sixth Annual Canada-Germany Symposium, in Montreal, Quebec, November 1995. I am grateful to my discussant, Harald Trabold of the Deutsches Institut für Wirtschaftsforschung, Berlin, for his insightful and gracious comments on the paper, as well as the conference participants, in particular Fanny S. Demers and Michel Demers, for a most stimulating discussion. Finally, I thank Murray G. Smith, former Director of CTPL, for inviting me to participate in the conference and for several valuable discussions on the subject matter of the paper and related topics; and Isolda Guevara of GTPL for her generous and timely logistical and administrative assistance prior to, during, and after the conference. The responsibility for all errors and opinions is my own.

The author is Assistant Professor of International Affairs, Carleton University, and has been a Visiting Professor of Economics at the Institute for Advanced Studies.
INTERNATIONAL TRADE AND THE DOMESTIC WAGE STRUCTURE

Alternative Explanations

1. Introduction

An important phenomenon that has recently been extensively documented by labour economists is the apparent growing divergence in the wage structure in the United States in the 1980s. While the wage structure is diverging along many dimensions, the one dimension that is of interest here is the widening skill differential, defined conventionally in the labour economics literature as the ratio of the wage of a college graduate over that of a high school graduate. These developments are well documented, in a lucid and elegant summary and extension of the literature by Jacob Mincer (1991), in a recent symposium issue of the Quarterly Journal of Economics (February 1992) and in a lead article in the American Economic Review, John Bound and George Johnson (1992), as well as in the lay press, e.g., the New York Times, Peter Passell (1992). More recent developments are explored in a colloquium issue of the New York Federal Reserve’s Economic Policy Review, January 1995, in papers presented at a Brookings conference in February 1995, and in a review article by Gary Burtless (1995), forthcoming in the Journal of Economic Literature. For a critical survey and synthesis of the literature, see Dehejia (1992a,b), Bhagwati and Dehejia (1994), and Bhagwati (1995). The chief finding of the literature is that the skill differential, after rising modestly in the 1950s and 1960s and actually falling in the 1970s, has risen sharply again in the 1980s. In particular, the skill differential has risen over 15 per cent from 1979 to 1988. The existence of this increase in the skill differential in the 1980s will be taken as a prior in this research. It is concisely displayed in Figure I, panel b, Katz and Murphy (1992), 42.

In the search for causes of this increasing skill differential, international trade is often mentioned as an important cause, by way either of a fall in the relative price of manufactured goods, skill-biased technical progress in the global economy, immigration (legal and illegal) of unskilled labour, and the large US trade deficit (see Borjas, Freeman, and Katz, 1992, Murphy and Welch, 1991, and Revenga, 1992). Unfortunately, with few exceptions (notably Grossman, 1986, 1987) the early labour economics studies of the effect of international trade on the wage structure—which claim to find a significant effect—are vitiated by a failure carefully to consider the general equilibrium effects of
international shocks and by a tendency to confuse quantity and price effects.\textsuperscript{16}

More recently, international trade economists have explored the causal relationship between international trade and the domestic wage structure. The evidence is mixed. Leamer (1995) presents evidence suggesting that trade has been an important cause of the widening wage differential and falling wages of the unskilled, whereas Lawrence (1995) argues to the contrary.

Further, although there is a plethora of empirical analysis of the issue, there is a paucity of theoretical trade models which attempt to tackle the issue squarely. Analysis typically revolves around the central propositions of neoclassical trade theory on income distribution, namely, the Stolper-Samuelson result and the Findlay-Grubert (1959) analysis of the effects of different assumptions on technical progress within the neoclassical model. As for theoretical models which directly address the issue, there is a paper by Ronald Findlay (1991) which elegantly extends the earlier Findlay-Grubert (1959) analysis to a 3x3 model and considers a variety of comparative statics exercises and more recent analyses by George Johnson and Frank Stafford (1992, 1995), but that to my knowledge is the extent of the theoretical literature. Ronald Jones (1994) provides a useful taxonomy of the trade-theoretic results that bear on the issue.

It would be germane briefly to summarise the chief findings of the neoclassical literature. The celebrated Stolper-Samuelson result is the chief construct, since it demonstrates that, in the context of the neoclassical 2x2 model, a change in the terms of trade improves the real wage of the factor which is intensively used in the good whose relative price has gone up and worsens the real wage of the other factor, assuming that the economy remains diversified and ruling out factor-intensity reversals. Labour economists who argue that a fall in the relative price of manufactures hurts the real wage of unskilled labour evidently have the Stolper-Samuelson theorem in mind. The second key result comes from the Findlay-Grubert (1959) analysis and demonstrates that Hicks-neutral technical progress in one sector will increase the real wage of the factor intensively used in that sector and decrease the real wage of the other factor. Thus, labour economists who argue that technical change in skill-intensive sectors is driving the fall in the wage of unskilled labour evidently have the Findlay-Grubert (1959) result in mind. The problem is that the evidence to my knowledge does not appear to show, for the US, a secular

\textsuperscript{16} Indeed, it is no surprise that the careful studies by Grossman, which model the price and quantity effects carefully and are alive to general equilibrium considerations, finds little effect of import competition on wages through the conventional Stolper-Samuelson channel.
decline in the relative price of manufactures (a secular increase in the terms of trade assuming unskilled-intensive manufactures to be US imports); protection, in the form of tariff and non-tariff barriers, including VERs, appear to have propped up the relative price of imported manufactures. A quick glance at, for instance, Figure 7 in Bhagwati (1991) verifies this intuition. Equally problematic is the assumption of exogenous technical change which favours skilled over unskilled labour; in light of modern theories of economic growth which attempt to endogenise technical change, the obvious question is, where does this technical change come from? In the absence of a theory of why technical change is skill-biased, a mere assertion that it is appears to me to be an unsatisfactory explanation.

In Bhagwati and Dehejia (1994), we explored a variety of hypotheses that have been proposed to account for the phenomenon of the rising skill differential, including the Stolper-Samuelson explanation and other trade- and non-trade-based explanations. Two of the models presented in that paper were a model of kaleidoscopic comparative advantage and capital-skill complementarity. The nature and scope of the paper necessitated a somewhat cursory treatment of the analytical underpinnings of the respective models, and that is the gap which this paper fills.

2. Kaleidoscopic Comparative Advantage

The approach that will be followed in this analysis is to build a model which attempts to fit the stylised facts of a widening skill differential in the context of a trading economy without making the assumption either of a secular decline in the relative price of manufactures or arbitrarily assuming skill-biased technical change. The approach will base itself, however, squarely on the 2x2 neoclassical model, since the rival explanations evidently use this model as a starting point. As will be discussed in the proceeding section, the approach will be to embed within the neoclassical model an assumption on human capital accumulation which is based on Jagdish Bhagwati’s ‘rolling stone’ hypothesis (Bhagwati, 1991)—more on this below. It is shown that augmenting the neoclassical model with ‘rolling stone’ delivers the outcome of a widening skill differential without assuming a secular improvement in the terms of trade—which is empirically false—and without merely assuming skill-biased technical change—which is arbitrary.

The essential insight of the ‘rolling stone’ hypothesis is summarised by Bhagwati (1991) as follows: “The tremendous globalization of investment and production in the world economy, accelerating in the 1980s, and the flexibility of exchange rates, have meant that the competitive advantage of specific industries and firms can be eroded
quickly through foreign competition. This probably creates greater flux and disruption in employment: increasing the forced mobility from job to job as jobs disappear and arise with greater frequency. If so, the effect on the earnings of HS [high school graduates] could be deleterious if these earnings reflect skills acquired on the job and often specific to the job. The higher turnover of jobs help would lead to repeated loss of skills appropriate to vacated jobs and also to loss of scale economies in learning over longer periods at any given job. Insofar as earnings reflect the acquisition of skills, there would be a corresponding adverse effect on earnings. The same adverse effect on earnings would not follow for the College graduates who, comparatively speaking, would enjoy returns to skills more from their education and less from on-the-job training” (13). In section 3 below I present a model which attempts to give structure to the verbal argument that Bhagwati exposit and which was further elaborated in Bhagwati and Dehejia (1994).

As for empirical evidence of the importance of job turnover, a study by the Organisation for Economic Cooperation and Development, cited by the Economist (1993), articulates a concern about the erosion of worker skills due to increased job turnover and provides evidence giving credence to this concern.

3. A Simple Model

Consider the 2x2 neoclassical model of international trade for a small country taking the terms of trade as given, as exposit for instance in Jagdish N. Bhagwati and T.N. Srinivasan (1983, Appendix B). Suppose that the two factors of production are highly-skilled labour, H, and less-skilled labour, L, where H and L denote the effective stocks of these factors. Normalising the populations of H-type and L-type workers to unity for simplicity, H and L are interpreted as the endowment of skills per highly-skilled and less-skilled worker, respectively. Suppose that the effective endowment of highly-skilled labour augments due to learning-by-doing at the constant and exogenous rate $\delta_H$, where the learning process of an H-type is not impeded by moving between the two sectors. By contrast, the effective endowment of less-skilled labour is assumed to augment at the rate $\delta_L$, but only for those workers who at any instant of (continuous)
time remain within a given sector; L-type workers who at any instant shift between sectors are assumed instantaneously to lose the benefits of learning. These assumptions, embodying Bhagwati’s (1991) ‘rolling stone’ hypothesis (see also Peter Passell, 1992 and Bhagwati and Dehejia, 1994) imply the following formulation for the dynamic evolution of the effective endowment of relative stocks of labour,

\[ \frac{dh}{h} = \delta_H - \delta_L - \delta_L |d\lambda|, \]

where the vertical bars denote the absolute value operator, \( d \) denotes the differential operator, where \( h \equiv H/L \) and \( \lambda \equiv L_1/L \), and where \( h \) is the relative endowment of highly-skilled labour and \( \lambda \) is the share of less-skilled labour employed in sector 1.\(^{18}\)

I also use the following important identity,

\[ h = h_1(\omega)\lambda + h_2(\omega)(1 - \lambda), \]

where the functions \( h_i(\omega), i = 1,2 \), denote the choices made by a cost-minimising entrepreneur in sector \( i \) at a given relative wage \( \omega \equiv w_L/w_H \). Writing (2) in differential form gives

\[ d\lambda = \frac{1}{h_1(\omega) - h_2(\omega)} dh - \frac{h_1(\omega)\lambda + h_2(\omega)(1 - \lambda)}{h_1(\omega) - h_2(\omega)} d\omega, \]

where it is assumed without loss of generality that \( h_1(\omega) > h_2(\omega) \) for all \( \omega \) over the range of interest, i.e., sector 1 is intensive in the use of highly-skilled labour. Recall that the first term in (3) captures the Rybczynski effect and the second term captures the Stolper-Samuelson effect.

Finally, the equation which relates the relative price of good 2, \( p \), to the relative wage, \( \omega \), is given by

\[ p = \frac{f_1(h_1(\omega))}{g_1(h_2(\omega))} = \frac{(f - h_1f_1)}{(g - h_2g_1)} \equiv c(\omega), \]

\(^{18}\) Alan Deardorff has suggested to me that the H- and L-types are better interpreted as “educated” and “skilled” workers, respectively, as the truly unskilled by definition have no skills that can become obsolete; by contrast, “skilled” workers, who train either in trade schools or on the job, are more susceptible to the erosion of their skills through turnover than “educated” workers, whose training comes from college or university settings and whose skills are relatively portable. However, for reasons of clarity, I will continue to use the terms “high-skilled” and “less-skilled”.
where \( f \) and \( g \) are the average product functions for sectors 1 and 2, respectively, and where subscripts denote partial derivatives. In differential form, this becomes

\[
dp = c'(\omega) d\omega. \tag{23}
\]

Finally, suppose that the terms of trade for this small country evolve according to a continuous time Gaussian stochastic process \( dp \) with zero mean and positive finite variance \( \sigma^2 \). This specification captures our assumption that the terms of trade are assumed to have no trend but fluctuate noisily around some (unspecified) constant long-run level due to (unspecified) world-wide technology or preference shocks. This is absolutely crucial; with a nonzero trend, the action comes from a Stolper-Samuelson effect; with zero trend we have pure white noise and the action comes purely from the ‘rolling stone’ effect.

Putting all of these elements together, it is possible to derive the following expression for the evolution of the growth rate of \( h \), the relative endowment of highly-skilled labour:

\[
\frac{dh}{h} = \delta_H - \delta_L - \delta_L \left( \frac{d\sqrt{h'}(c^{-1}(p))\lambda + h'_2(c^{-1}(p))(1 - \lambda)dp}{h_1(c^{-1}(p)) - h_2(c^{-1}(p))} \right), \tag{24}
\]

where equation (4) has been inverted and been used to substitute out for \( \omega \), using Samuelson’s univalence result on the monotonicity of \( c(\omega) \).

Now, the ‘skill differential’ or ‘return to education’ in this model is evidently

\[
h/\omega \equiv (H/L)(w_H/w_L),
\]

---

19 It is straightforward to extend this model to the large country case by making fluctuations in the terms of trade depend in turn on more primitive shocks, such as for instance preference or technology shocks, but this makes no qualitative difference to the analysis.

20 It is of course possible to construct a hybrid formulation which incorporates a trend effect as well, something not attempted here.

21 This formulation will hold strictly so long as the economy remains within the Chipman-McKenzie diversification cone; if it falls outside the cone and specialises on either commodity, turnover of labour will cease and therefore H- and L-types will augment their human capital at the rates \( \delta_H \) and \( \delta_L \) respectively with a permanent level difference that was accumulated during the differential augmentation within the diversification cone.
which is the ratio of the real-wage of an H-type to the real wage of an L-type. Hence, we are interested not only in \( dh/h \), which tracks the evolution of relative effective endowments, but \( dh/h - d\omega/\omega \), which tracks the evolution of the return to equation. Of course, this being a stochastic model, what really interests us is not the actual paths of these entities—which cannot in general be determined—but their expected paths—which in principle can be.

Unfortunately, due to the inherent nonlinearity of the absolute value operator, equation (6) is not amenable to analytical solution. However, the following conjectures on the possible path of \( h \) are offered. Firstly, the expected path of \( h \) presumably depends upon the gap between the exogenous learning rates of highly- and less-skilled workers, \( (\delta_H - \delta_L) \), so that \( h \) will augment more rapidly the bigger the deviation. This is quite obvious, but somewhat less obviously it presumably depends upon the standard deviation, \( \sigma \), of the white noise disturbance to the terms of trade, i.e., I speculate that, the noisier are the terms of trade, the greater is the rate at which human capital accumulation by highly-skilled workers outpaces the accumulation of less-skilled workers, in other words, the greater the rate at which real incomes of highly-skilled workers diverge from the real incomes of less-skilled workers. If \( \sigma = 0 \), so that the terms of trade are noise-free, and if the ‘natural’ rates of learning \( \delta_H \) and \( \delta_L \) are equal, the expected path of \( h \) will depend only on the discrepancy between the exogenous learning rates, as is conventional. Furthermore, since \( h/\omega \) is the product of a variable with a rising trend, \( h \), and one which is purely noise, \( \omega \), \( h/\omega \) will presumably evolve noisily around a rising trend.

To buttress these conjectures, a number of simulation exercises of the model have been implemented. Evidently, to simulate a model of the type presented here, it is essential to specify functional forms for the generalised functions used above and select parameter values. I select therefore Cobb-Douglas functional forms for the technology, with parameter on \( H \) being \( \alpha \) for good 1 and \( \beta \) for good 2. Then, it is possible to solve for the functions \( h_i(\omega) \). It turns out that \( h_1(\omega) = (\alpha/(1 - \alpha))\omega \) and similarly for \( h_2(\omega) \). These in turn determine \( c(\omega) \). It is then necessary to select starting values for \( H \) and \( L \) and specify the stochastic process for \( dp \). I selected a process with stationary white noise disturbances with parameter \( \sigma \). \( \sigma \) measures the noisiness of the terms of trade. In the simulations presented, \( \alpha \) was selected to equal 0.8, \( \beta \) to equal 0.6, \( \delta_H \) and \( \delta_L \) were both set equal to 0.01 as a benchmark case, and starting values for \( H \) and \( L \) were set at 2.5 and 1 respectively.\(^{22}\) In the simulations presented, the noise parameter,

\(^{22}\) These values were chosen in part to ensure that the economy remained within the
σ, is successively set equal to 0.1, 0.5, and 1. Plots of the paths of \( h/\omega \), \( h \), and \( \omega \) under the three values for \( \sigma \) are presented in Figures 1, 2, and 3, respectively.\(^{23}\) Examination of the three figures bear out some of the conjectures presented above. A higher \( \sigma \), which results in a successively noisier \( \omega \), is shown successively to imply a steeper accumulation of \( h \).\(^{24}\)

As one would expect, \( h/\omega \) inherits both the trend and the noise from \( h \) and \( \omega \), respectively, but in Figures 1 and 3, with low and high noise respectively, the noise appears to overwhelm the trend for the sample presented. Only in the simulation with intermediate noise, Figure 2, is a clear trend discernable amidst the noisy fluctuations. My intuitive explanation for this phenomenon—which I did not initially expect—is that when noise is low, ‘rolling stone’ has little role to play, so that in the absence of a difference in natural learning rates, \( h \) will remain roughly constant and \( h/\omega \) will be dominated by the noise from \( \omega \); by contrast, when noise is high, ‘rolling stone’ is playing a significant role, so that \( h \) is distinctly rising sharply, but the very high noise from \( \omega \) is overwhelming the trend growth in \( h \) and therefore \( h/\omega \) appears to show little trend over the sample period. Certainly, glancing at all three figures, one is struck by the richness in the evolution of the ‘skill differential’; although it is generally rising in this model, the upward trend is quite often overwhelmed by a change in \( \omega \), so that, at least over a short period, an increase in the relative wage of unskilled labour (purely due to a white noise disturbance) lessens the skill differential, at least for that period.

The novelty of this model, apparent from the figures, lies precisely in the fact that the differential human capital augmentation and consequent divergence in real incomes between highly- and less-skilled workers arises solely due to the ‘rolling stone’ phenomenon. No Stolper-Samuelson-type trend in the terms of trade has been assumed - the terms of trade were allowed to fluctuate noisily but remained stationary around a long-run level - nor has there been any arbitrary assumption on differential rates of technical progress. The rising ‘skill differential’ is an endogenous phenomenon arising out of the interaction of a noisy terms of trade with the differential nature of human capital accumulation by skilled and unskilled workers.

---

\(^{23}\) Recall that \( \omega \) is a one-to-one map of \( p \) by Samuelson’s univalence theorem.

\(^{24}\) I also show but do not present here the result that in the absence of noise \( h \) remains stationary at its initial level when the natural rates of learning are equal.
4. Capital-Skill Complementarity

Recall from the discussion in section 1 that the labour economics evidence strongly appears to favor relative demand shifts as the proximate cause of the changing wage structure, with relative supply shifts moving in an offsetting direction but evidently of insufficient magnitude to prevent a widening of the wage structure. In this respect, the 1980s appear to contrast with previous periods of a widening wage structure which appear to have been followed by a relatively rapid relative supply response and return to a long-run level of the 'college premium' of 6 to 8 per cent. Thus, a second puzzle in the 1980s is, why has the supply response been so muted?²⁵

Amongst the plethora of possible causes of the relative demand shift explanation that have been proposed and tested, the consensus in the labor literature increasingly appears to favor skill-biased technical change, possibly related to the computer revolution, as a major contributing cause. This hypothesis has been put forward by Davis and Haltiwanger (1991), Krueger (1991), and Mincer (1991), and has recently been forcefully supported empirically by Bound and Johnson (1992). The basic argument is that the implementation of new technologies in the workplace, say robotics, makes it possible to substitute away from unskilled labor, but that these new technologies nonetheless require skilled personnel to operate them.

Such arguments are naturally cast in terms of capital-skill complementarity, viz., that new capital or new technology tends to be relatively complementary with skilled labor and relatively substitutable with unskilled labor, a factor which one would expect to lead to greater divergence in the wage structure.²⁶ Indeed, direct evidence of this complementarity has been presented by Griliches (1969) and more recently by Bartel and Lichtenberg (1987) and Berndt and Morrison (1991).

The purpose of this section of the paper is to present the simplest possible model of an endogenous wage structure which is based upon a minimal set of underlying assumptions, in particular, on the assumption of capital-skill complementarity--giving the relative demand effect--and furthermore based upon the assumption that the transformation of labor from an unskilled occupation to a skilled occupation is an inherently costly process and hence one which occurs slowly over time. Of course, whenever adjustment is costly and hence time-consuming, the adjustment problem of rational agents is

²⁵ This puzzle appears first to have been noticed by Jacob Mincer.
²⁶ It is not suggested that capital and skill are necessarily strictly complementary, but that they are less substitutable than capital and unskilled labor as measured by the Hicks-Allen partial elasticity of substitution.
inhertently dynamic and forward-looking and hence must be modelled akin to investment theory.

To anticipate the chief findings, it is demonstrated that, when capital (or technology) and skill are complementary with respect to unskilled labor, a positive shock to the capital stock will lead to an instantaneous jump in the skill differential and a gradual return to steady-state (characterized in this simple model by no differential in wages) accompanied by a fall in the skill differential as labor undertakes slow and costly retraining in response to the capital- or technology-induced relative demand shock. Furthermore, the instantaneous jump in the skill differential following the shock will be smaller in magnitude under rational expectations than under static expectations. Finally, the speed of adjustment and hence the speed with which a rise in the skill differential induced by a relative demand shock disappears is shown to depend crucially on the assumed cost of adjustment function in the economy's retraining sector; if retraining is extremely costly, adjustment will be slow and the skill differential will persist a long time. The model therefore offers a possible explanation of the relative supply puzzle alluded to above; if the model is correct, the tardiness in the relative supply response in the 1980s can be attributed to the fact that it was in some sense costlier for an unskilled worker in the 1980s to upgrade his skill level and occupy a skilled position that in previous decades. This explanation is consistent with the evidence, both formal and anecdotal, of declining standards of quality in the US educational structure, which (presumably) makes it costlier in real terms for an unskilled person to become skilled.

---

27 Similarly, it can be shown that a transitory shock leads to a smaller jump in the skill differential than a permanent shock of the same magnitude.
5. Setting up the model

Consider an infinitely-lived economy which produces a single output good, $Y$; with the assistance of three inputs, $K$, which represents 'capital' or 'knowledge', 28 $L_S$, skilled labor, and $L_U$, unskilled labor, according to a production function $F(K, L_S, L_U)$. The model will be a medium-run model in the sense that the stock of capital or knowledge $K$ will be considered exogenous and the total workforce, $\bar{L}$, will be taken as constant, so that $\bar{L} = L_S + L_U$. I would like to capture the idea that $K$ and $L_S$ are relatively complementary as compared to $L_U$, for which the following nested constant elasticity of substitution (CES) production function, suggested in Layard and Walters (1978), 275 - 276, proves useful:

$$Y = \{ \delta [\alpha K^{-\rho_1} + (1 - \alpha)L_S^{-\rho_1}/\rho_1 + (1 - \delta)L_U^{-\rho_2}]^{-1/\rho_2} \}, \quad (25)$$

where the condition that $\rho_1 > \rho_2$ guarantees that capital $K$ and skilled labor $L_S$ have a lower Hicks-Allen partial substitution elasticity than $K$ and unskilled labor $L_U$, so that capital-skill complementarity holds in a relative sense. The general results of this model hinge only on the assumption about the relative partial substitution elasticities, not on the specific functional form chosen; the functional form is presented to demonstrate the simplest possible production function that can capture the assumptions and may prove tractable for future empirical implementation of the model.

I suppose that the representative, perfectly competitive firm in the output-producing sector hires capital in a frictionless auction market at a rental price which I normalize to unity. By contrast, labor must be hired from a perfectly competitive 'retraining' or 'schooling' sector which purchases (sells) unskilled labor services at a price $q_U$ and sells (purchases) skilled labor services at a price $q_S$. The dynamic optimization problem of the representative output-producing firm is then to maximize the present discounted value of instantaneous cash flows; it therefore solves the following problem:

$$\max_{\{K, I_S, I_U, L_S, L_U\}} \int_t^\infty (Y - K - q_SI_S - q_UI_U)e^{-r(\tau-t)}d\tau \quad (26)$$

$$s.t. \dot{L}_S = I_S, \dot{L}_U = I_U \quad (27)$$

where $r$ is the constant and exogenous discount rate and where the initial split of the labor force into skilled and unskilled, $\bar{L} = (L_S)_t + (L_U)_t$, is assumed as given.

28 I am at this stage agnostic as to the precise interpretation of $K$; see the concluding discussion below.
The representative firm in the ‘retraining’ or ‘schooling’ sector is assumed to buy and sell unskilled and skilled labor services and is assumed able to transform unskilled into skilled labor and vice-versa according to an increasing, convex cost of adjustment technology. Assuming without loss of generality that the ‘retraining’ sector is retraining unskilled labor to be skilled, the dynamic optimization problem of the representative ‘retraining’ firm is

\[
\max_{\{I_S\}} \int_1^{\infty} ((qs - qu)I_S C(I_S)) e^{-r(t-t')} d\tau
\]

where it is assumed that \( C(I_S) \equiv q_S L_T(I_S) \) satisfies \( L_T(0) = L'_T(0) = 0 \) and \( L''_T(I_S) > 0 \forall I_S \). This specification of the adjustment technology makes the assumption that the ‘retraining’ sector employs skilled labor, which seems a reasonable first-order approximation of reality.\(^{29}\) The specification also closely follows the standard assumptions in neoclassical investment theory and ensures a smooth and continuous investment response to any differential in wages between sectors.\(^{30}\) Furthermore, the specification ensures that, for small disturbances around steady-state, I am entitled to ignore the resources employed in the ‘retraining’ sector since these are technically of a second-order of magnitude. It will prove useful at a later stage to specialize the cost of adjustment function to be quadratic,

\[
C(I_S) = \frac{q_S}{2\beta} I_S^2,
\]

where \( \beta > 0 \) is a parameter which captures the speed of adjustment and is inversely related to the cost of adjustment.

I am now ready to characterize the model’s equilibrium conditions. The first-order condition for the ‘retraining’ sector’s problem is:

\[
(q_s - qu) = C'(I_S),
\]

which equates the marginal return to retraining or schooling to its marginal cost; this first-order condition implies that the optimal rate of retraining is a strictly increasing

\(^{29}\) It is certainly the case that the largest single item in the budget of a university, for instance, is payroll expenditure.

\(^{30}\) The adjustment technology used here is similar to that used in Mussa (1978) and in Gavin (1990, 1991, 1992). When these assumptions are departed from, hysteresis becomes a possibility, which, while not uninteresting, is not the focus of the present paper. Insights on the hysteresis case in a related model may be found in Kemp and Wan (1974).
function of the difference between the price of labor in the two sectors:

\[ I_S = I(q_S - q_U) \equiv I(\Lambda), I \equiv C^{-1}, \]  

where \( I(\Lambda) \) denotes the inverse of the marginal cost of retraining function and where \( \Lambda \equiv (q_S - q_U) \) denotes the absolute differential in the prices of the two types of labor. In the special case of quadratic adjustment costs, this 'retraining' function is linear in \( \Lambda \):

\[ I_S = \frac{\beta}{q_S} \Lambda. \]  

This 'retraining' or 'schooling' function is analogous to Tobin's \( q \) theory of investment, except that what is being modelled is the flow of labor from one sector to another, which depends upon the difference between the two sectors' \( q \)'s.

Consider now the the first-order conditions for the output- producing firm. The following Euler equations for the firm are formally derived in appendix A.1:

\[ \frac{\dot{q}_S}{q_S} + \frac{w_S}{q_S} = r, \]  

\[ \frac{\dot{q}_U}{q_U} + \frac{w_U}{q_U} = r, \]  

where \( w_S \equiv F_{L_S} \) and \( w_U \equiv F_{L_U} \), and where subscripts denote partial derivatives and it is assumed that the derivatives are computed at the appropriate equilibrium values. These conditions set the total return on a unit of labor, including both 'capital gains' and the value of the marginal product of labor, equal to the opportunity cost of hiring labor, given by the constant and exogenous interest rate \( r \). Subtracting (15) from (14) yields:

\[ \dot{\Lambda} = r\Lambda - (w_S - w_U) \equiv r\Lambda - \omega, \]  

where \( \omega \equiv (w_S - w_U) \) and which can be integrated forward, imposing the appropriate transversality conditions, to give:

\[ \Lambda_t = \int_t^\infty \omega e^{-r(\tau-t)} d\tau. \]  

I obtain the completely intuitive result that the difference in the prices of the two types of labor will equal the present value of expected future differences in wages between the two types of labor, highlighting the forward-looking nature of the problem.

It is important to note that \( \omega \) will be a monotonically declining function of \( L_S \), the total 'skilled' workforce, and will be a monotonically increasing function of \( K \), the
aggregate stock of 'capital' or 'knowledge'. The first fact is standard and follows from diminishing returns - as more people become skilled, the absolute and relative rewards to being skilled go down - and the second fact captures the capital-skill complementarity hypothesis and follows directly from the assumptions on the Hicks-Allen partial substitution elasticities.\(^\text{31}\)

In the special case of the production function in equation (7), \(\omega\) is shown to be given by:

\[
\omega(K, L_S) = \frac{-1}{\rho_2} \{\delta[\alpha K^{-\rho_1} + (1 - \alpha)L_S^{-\rho_1}]^{-\rho_2/\rho_1} + (1 - \delta)(\bar{L} - L_S)^{-\rho_2} \}^{-1 - \rho_2/\rho_2} \\
x\{-\delta \rho_2 (1 - \alpha)(\alpha K^{-\rho_1} + (1 - \alpha)L_S^{-\rho_1})^{-\rho_2-\rho_1/\rho_1} L_S^{-\rho_1-1} - (1 - \delta) \rho_2 (\bar{L} - L_S)^{\rho_2-1}\}
\]

where it is easily verified that \(\omega_K > 0\) and \(\omega_{L_S} < 0\).

\(^{31}\) Indeed, the model structure is very general; any underlying production structure satisfying the two conditions on the partial derivatives of \(\omega\) will yield the same qualitative conclusions, including, for instance, a two-sector Ricardo-Viner economy in which the sector employing skilled labor is relatively \(K\)-intensive as compared to the sector employing unskilled labor.
6. Steady-State equilibrium and transitional dynamics

The model is characterized by an autonomous system of two differential equations, (12) and (16), in two variables, \( L_s \) and \( \Lambda \). Steady-state in the model is characterized by the following conditions:

\[
\Lambda = 0, \Delta r = \omega,
\]  

which implicitly determines a steady-state \( L^*_s = \omega^{-1}(0) \) and hence an equilibrium split of the labor force between the 'skilled' and 'unskilled' at which their wages are equalized. Notice that, under the maintained assumptions, \( L^*_s \) is increasing in \( K \) since \( \omega \) is increasing in \( K \). This is eminently intuitive: since, for a given split of the workforce between skilled and unskilled, a jump in \( K \) leads to an increase in the relative reward to skill, in steady-state in which there is no incentive for an unskilled person to become skilled it must be the case that a larger proportion of the workforce is skilled as compared to before the jump in \( K \). This establishes immediately that the capital-skill complementarity hypothesis cannot, in this model, have steady-state implications for the wage structure, since steady-state is inconsistent with any difference in the wages of skilled and unskilled.\(^{32}\)

Considerable insight on the adjustment process may be gained by examining the transitional dynamics of the system.\(^{33}\) For this purpose, consider figure 4, which maps the steady-state loci in the \((\Lambda, L_s)\) phase plane. The \( \hat{L}_s = I_s = 0 \) steady-state condition, plotted as the II locus, is coincident with the horizontal axis (at which \( \Lambda = 0 \)). The \( \hat{\Lambda} = 0 \) steady-state condition, plotted as the \( \Lambda \Lambda \) locus, is downward-sloping in the plane since \( \omega \) is monotonically declining in \( L_s \), as discussed.

\( L^*_s \) is thus the point at which the downward-sloping \( \Lambda \Lambda \) locus crosses the horizontal axis. Qualitative dynamics in the four regions of the phase plane are indicated by the arrows. It is intuitively evident (and is formally proved in appendix A.2) that there exists a downward-sloping saddle path in the plane, labelled \( SS \), which furthermore will

\(^{32}\) In this simple model, there is no steady-state wage differential to compensate skilled workers for the cost of becoming skilled because of the assumption that the costs of adjustment are negligible for small deviations around steady-state. The model could certainly be modified to take account of a difference in steady-state wages, an extension not attempted here.

\(^{33}\) This analysis is purely a local analysis for small disturbances around the steady-state, i.e., what is considered are the linearized dynamics around the steady-state. Global analysis, while of interest, is not part of this paper's emphasis.
be flatter than the $\Lambda$ locus and along which the approach to the steady-state will be monotonic. Equilibrium requires that, given an initial split of the labor force and hence an initial $(L_S)_t$, $\Lambda$ be chosen to equal $\Lambda_t$ so that the economy jumps onto the saddle path $SS$; any other initial choice of $\Lambda$ will put the economy onto an aberrant trajectory which will eventually lead to a violation of the transversality conditions.

Consider now the effect of an exogenous and permanent increase in $K$, assuming that the economy begins from a previous steady-state. As I have demonstrated earlier, an increase in $K$ increases $\omega$, which implies an upward shift in the $\Lambda\Lambda$ locus, hence an increase in steady-state $L_S^*$ and an upward shift in the saddle-path $SS$ to pass through the new steady-state. To avoid cluttering the diagram, suppose that $(L_S)_t$ was the steady-state coinciding with the old $\Lambda\Lambda$ and $SS$ loci (undrawn) and that the new steady-state is at $L_S^*$. Then, adjustment to the shock will be exactly as described above: in the immediate aftermath of the shock, there will a jump in the relative reward to skilled labor of just sufficient magnitude to put the economy on a stable trajectory along which unskilled workers gradually retrain and become skilled and along which the ‘skill differential’ is thus slowly bid away by the endogenous supply response.

Two notes on the adjustment process are in order. First, the adjustment process under perfect foresight is less rapid than under static expectations. To see this, notice, following Lucas (1966) and Mussa (1978), that, under static expectations, the economy must always travel toward steady-state along the $\Lambda\Lambda$ locus - since by definition of static expectations no change in $\Lambda$ is expected. But then, at every value of $L_S$, except $L_S^*$, the absolute value of $\Lambda$ under static expectations is larger than the absolute value of $\Lambda$ under rational expectations; thus, the amount of labor devoted to retraining and the equilibrium $q_S$ must always be higher under static than rational expectations. Therefore, the economy converges to its steady-state equilibrium faster under static than rational expectations. Intuitively, the incentive to adjust is smaller and hence the adjustment process is slower when forward-looking agents recognize that the current wage differential will be falling in the future than when they myopically expect the current differential to persist indefinitely. However, more rapid adjustment would not be desirable, since the perfect foresight adjustment path maximizes the discounted value of the economy's output (proven in appendix A.1).

Second, the adjustment process along the perfect foresight adjustment path depends crucially on the cost of adjustment technology. To see this most clearly, consider the ‘retraining’ function in the special case of quadratic adjustment costs, given by equation (13). For any given $\Lambda$, the investment response depends linearly on $\beta$, the speed of
adjustment parameter of the cost function. For a given $\Lambda$, as $\beta$ goes to zero, adjustment becomes increasingly costly and hence increasingly slow. In the extreme case of $\beta = 0$, adjustment is infinitely costly and no adjustment ever takes place: the division of the workforce between skilled and unskilled is given by history and inequality will persist indefinitely. Conversely, as $\beta$ goes to infinity, adjustment becomes less costly and hence increasingly rapid; in the limit, when $\beta$ is set equal to infinity, adjustment is instantaneous; $\Lambda$ can never deviate from zero since even the slightest deviation would lead to potentially infinitely large flows of labor between skilled and unskilled occupations. In such a universe, perfect equality exists at all times and is impervious to any shocks. In light of this model, the observation that adjustment to a wage differential in the 1980s has been slow compared to the response to previous shocks would be interpreted by arguing that (for some reason) adjustment in the 1980s is costlier, say because of the declining quality of schools, and hence adjustment is proceeding more slowly than before.34

7. Discussion

This paper has presented two models which address two alternative explanations of the wage differential phenomenon. The first model, the ‘kaleidoscopic comparative advantage’ or ‘rolling stone’ model, was designed to generate a widening skill differential without assuming a secular trend in the terms of trade à la Stolper-Samuelson. It did this in the context of the neoclassical 2x2 model for international trade by simply augmenting it with a dynamic equation relating to the relative acquisition of human capital by skilled and unskilled workers as a function of labor turnover. Numerical simulations of the model confirmed the conjecture that greater volatility in the terms of trade induce a more rapid divergence in the wage structure between skilled and unskilled workers. The chief remaining issue here is empirical, viz., how relevant is the hypothesis? This is left as the subject of future research.

The second model, the ‘capital-skill complementarity’ model, attempted to rationalize in the most basic way possible the insight, drawn from the labor literature, that complementarity between skilled labor and capital (or knowledge) can account in large measure for the wide discrepancy observed in the 1980s between the wages of skilled and unskilled labor. However, as the model emphasizes, any discrepancy between wages will induce in forward-looking workers a process of rational adjustment as the unskilled ‘go back to school’ or retrain to become skilled, and hence, in the steady-state, this

34 More on this below.
wage differential is bid away due to the endogenous relative supply response of workers. Therefore, a clear prediction of the model presented in this paper is that the large discrepancy in wages between skill levels observed in the 1980s is a transitory phenomenon. Indeed, it is precisely the difference in wages which sets up the endogenous supply response which will eliminate that very wage difference in the long-run, suggesting that policy intervention, say in the form of a wage subsidy to the unskilled, aimed at closing the wage gap would be misguided and counterproductive, as it would only slow the adjustment process needed to restore long-run equilibrium. This insight seems to me to be fairly robust to the niceties of the modelling structure.

What about policy intervention designed to speed the adjustment process, say a subsidy designed to lower the cost of retraining and hence make it cheaper for an unskilled worker to become skilled? On this count, the model as it stands cannot seriously be used to offer policy advice, since, by construction it has the feature that the adjustment path is efficient, given the cost of adjustment technology and the assumption of perfect foresight: the costlier is adjustment, the slower is the perfect foresight efficient adjustment path. However, it would be facile to conclude therefore that the adjustment path is in some *ex-ante* sense efficient, albeit being *ex-post* efficient. While it would be difficult sensibly to argue for intervention based on a deviation from rational expectations, an important limitation to this (and similar) models is that the adjustment technology is taken as a primitive; if, however, we imagine that the adjustment technology itself depends upon some ‘deeper’ economic structure which has been left outside the model, we presumably might be less sanguine about concluding that the adjustment path is necessarily efficient, in the absence of plausible arguments that the adjustment technology itself has been (in some sense) efficiently chosen.\(^{35}\) Indeed, I am planning an extension of this model in which the adjustment technology is endogenized and hence in which the policy question can be more reasonably answered. This extension of the model would also allow a deeper answer to the question, why has adjustment been so slow in the 1980s? While I have suggested above that this model’s answer would be couched in terms of costlier adjustment in the 1980s, a deeper and more policy-rich

---

\(^{35}\) A model which goes part of the way towards addressing these concerns is presented in Grossman and Shapiro (1982), in which it is supposed that workers *ex-ante* choose the degree of sectoral mobility that they would like based on an expectation of the degree of economic volatility expected in the following period. Their argument, however, depends crucially on the presence of uninsurable uncertainty, whereas uncertainty is completely absent from this perfect foresight model.
answer would surely explain *why* adjustment is costlier; in the absence of such a deeper explanation, the answer remains somewhat shallow.

Finally, the model as it stands depicts a universe in which output is stationary in the steady-state, which is obviously countefactual. This shortcoming could be remedied by marrying the model with a standard model of growth, in which the evolution of $K$ would be endogenized and in which total population, $\bar{L}$, would also be allowed to grow. Such a growth-theoretic extension is also the subject of future research, and in addition it will allow me to draw a sensible distinction between ‘capital’ and ‘knowledge’ which can innocuously remain nebulous in this medium-run model.
APPENDIX

A.1

Consider the dynamic optimization problem of a representative output-producing firm maximizing (8) subject to (9) and standard transversality and ‘no bubble’ conditions. The current-value Hamiltonian to the problem is given by:

$$H = F(K, L_S, L_U) - K - q_S I_S + \lambda_S I_S - q_U I_U + \lambda_U I_U,$$

where $\lambda_S$ and $\lambda_U$ are the current-value costate variables corresponding to the state variables $L_S$ and $L_U$ respectively. The first-order conditions for a maximum are given by:

$$F_K = 1,$$

$$q_S = \lambda_S,$$

$$q_U = \lambda_U,$$

$$\dot{\lambda}_S = r \lambda_S - H_{L_S} = r \lambda_S - F_{L_S},$$

$$\dot{\lambda}_U = r \lambda_U - H_{L_U} = r \lambda_U - F_{L_U},$$

where equation (A.2) sets the marginal product of capital equal to its rental price (unity), equations (A.3) and (A.4) set the marginal cost of an additional increment to the stock of skilled and unskilled labor equal to the shadow value of an additional increment to the stock of skilled and unskilled labor, respectively, and equations (A.5) and (A.6) are akin to the usual asset-pricing equations. Combining (A.3) and (A.5) and (A.4) and (A.6), we obtain equations (8) and (9) respectively in the text. As in the text, we subtract (9) from (8) to obtain equation (10) in the text. We showed in the text that the rate of labor flow between occupations is given by equation (6). But since equations (6) and (10) implicitly maximize the discounted value of the representative output-producing firm’s cash flows (and the representative ‘retraining’ firm uses negligible resources in the vicinity of steady-state), this proves that the perfect foresight adjustment path defined by equations (6) and (10) is efficient, in the sense that maximizing the discounted value of a representative output-producing firm’s profits is equivalent to maximizing the discounted value of the economy’s final output in this perfectly competitive, undistorted economy.
A.2

Linearizing (6) and (10) around a steady-state, we obtain the following system,

\[
\begin{bmatrix}
\dot{\Lambda} \\
\dot{L}_S
\end{bmatrix} =
\begin{bmatrix}
r & -\omega_{L_S}(\cdot) \\
I'(\cdot) & 0
\end{bmatrix}
\begin{bmatrix}
\Lambda \\
L_S - L_S^*
\end{bmatrix},
\]

(A.7)

where it is understood that the partial derivatives are evaluated at the steady-state. The characteristic roots of the system are given by:

\[
\gamma_1 = \frac{r - \sqrt{r^2 - 4\omega_{L_S}(\cdot)I'(\cdot)}}{2}, \quad \gamma_2 = \frac{r + \sqrt{r^2 - 4\omega_{L_S}(\cdot)I'(\cdot)}}{2}.
\]

(A.8)

Recalling that \(\omega_{L_S}(\cdot)\) is negative and \(I'(\cdot)\) is positive, we see that \(\gamma_1\) is negative, while \(\gamma_2\) is positive, thus establishing the saddle-path stability property that I appealed to in the text. The eigenvector associated with the stable eigenvalue is

\[
\begin{bmatrix}
1 \\
e_1
\end{bmatrix} = \begin{bmatrix}
1 \\
I'(\cdot)/\gamma_1
\end{bmatrix}
\]

(A.9)

and since the second element element of the eigenvector is negative, we know that the saddle-path is downward-sloping in the \((\Lambda, L_S)\) plane as claimed in the text.
REFERENCES


Findlay, Ronald and Harry Grubert, 1959, Factor Intensities, Technological Progress and the Terms of Trade, Oxford Economic Papers, February, 111 - 121.


Economy 90, 1054 - 1069.

Johnson, George E. and Frank P. Stafford, 1992, Models of International Competition and Real Wages, mimeo (University of Michigan, Ann Arbor, MI).


Jones, Ronald W., 1994, Trade and Wages: Comments on Theoretical Issues, mimeo (University of Rochester, Rochester, NY).


Institut für Höhere Studien
Institute for Advanced Studies
Stumpergasse 56
A-1060 Vienna
Austria

Phone: +43-1-599 91-145
Fax: +43-1-599 91-163