Dynamic skills acquisition choice – Jacks of all trades vs. dab hands*

A. Sarychev†

December 28, 1999

Abstract

This paper proposes a dynamic theory of adjustment of labor force to a shock that makes existing human capital stock obsolete. Workers will not invest in the new, superior skills because their current value depends on the existing complementary stock of human capital, and the obsolete specialized skills are largely incompatible with the new specialized skills. Labor market imperfections make the current skills distribution an important factor in the decision to invest in new skills. This creates room for generalists, workers with an intermediate set of skills who are able to work with both old and new types. Along the equilibrium path the economy accumulates a buffer stock of generalists that eventually makes it profitable to invest in superior specialization. The paper proposes new methods of studying the dynamics of adjustment in search models with forward-looking investment. It characterizes the dynamics of transition and analyzes how equilibrium paths differ across countries with diverse labor market and educational institutions. The efficiency analysis allows drawing policy implications. Econometric evidence on labor markets in transition economies is shown to be broadly consistent with predictions of the model. East Germany presents a stark example of rapid transition which is difficult to explain by traditional theories but is consistent with the predictions of my model.

1 Introduction

Recently, a number of papers contributed to the analysis of equilibria in a general class of search models with heterogeneous workers, see for example Shimer and Smith [2000], Burdett and Coles [1997], and Lu and McAfee [1996]. However, there is a certain caveat in that literature, regarding the out-of-steady-state adjustment dynamics. At the same time there are real-life contexts, in which what happens during this adjustment is extremely important. This paper proposes a dynamic theory of adjustment of labor force to a shock that makes

*I benefitted from suggestions and advise of Daron Acemoglu, Abhijit Banerjee, Olivier Blanchard, Michael Kremer, and Sendhil Mullainathan, to whom I am expressing my sincere gratitude. I thank Jörn-Steffen Pischke for allowing me access to the GSOEP data.

†e-mail: dupel@mit.edu

existing human capital stock obsolete, and relates this theory to the recent developments in formerly-planned economies.

At the end of the eighties, a number of Central and Eastern European countries embarked on transition from plan to market. The liberalization shock shook everything: contractual arrangements available to firms, the structure of relative prices, and production technologies, thereby dramatically affecting relative rewards to different sets of skills. Skills inherited from the socialist era were not particularly attuned to the needs of market economy and so have been rapidly devaluing. Another salient pattern of the stock of human capital was its excessive specialization compared to most Western economies. The World Bank Development Report [1996] documents that education systems in planned economies have emphasized awareness of facts, in contrast to the stress on the *use* of knowledge in unanticipated circumstances in market economies. For example, in Poland, secondary technical schools had separate curricula for 300 different occupations, compared to 16 broad apprenticeship programs in West Germany. Furthermore, adult education and training in command economies was neglected. Consequently, labor force was characterized by much narrower specialization, as it conformed to demands of a largely static production structure.

Thus, workers were excessively specialized in wrong things. The expectation was that over time the structure of human capital in CEE countries would gradually converge to that in the West. The transition could only be gradual because skills are embodied in human beings and cannot be instantly upgraded. The question that received relatively little attention in the literature is how new skills interact with the old skills inherited from the planned economy *during the transition*.

In the long run, all workers will specialize in the *right* things (efficient grade of human capital). However, in labor markets characterized by costly mobility and search, new entrants into the labor force are ex ante uncertain about their future matches. They may not benefit from investing in the set of new special skills, if these skills are incompatible with the (obsolete and excessively specialized) skills of the majority of labor force. Individual-specific uncertainty about the match creates a pecuniary externality that discourages efficient investment and makes room for *general* human capital. Workers who are flexible enough to be both able to work with the obsolete specialists and to take advantage of the new productive techniques, will do well in the short run; they will also do relatively well when the efficient specialists eventually appear. Hence, general human capital may be more attractive at the beginning of transition until at some point there are enough generalists to make specialization pay off. The moment when specialization takes off is central to the nature and speed of adjustment. Its timing affects output growth, evolution of unemployment across groups, and inequality.

This paper presents a model in which workers may choose among a few grades of human capital that have differential productivities and are not all compatible with each other. Given a frictional labor market, expected earnings depend on the future evolution of the aggregate skills distribution. This makes human capital acquisition decisions interdependent and characterized by inertia, as agents take into account current and expected labor market conditions. Evolution of the skills distribution is in turn determined by those acquisition decisions.

In this framework I analyze implications of an aggregate shock that enlarges the set of possible human capital grades available to the economy. To deal with complex adjustment dynamics somewhat neglected by existing research, I introduce new analytical methods. I analyze explicitly the stochastic process underlying labor market transitions by studying its generator matrix. This approach allows me to derive analytical results; relying solely on the properties of that matrix I am able to characterize the relevant policy thresholds. I then study how they depend on the parameters of the model, and how they affect trajectories of unemployment, output per capita, and wage inequality. Application of a technique similar to calculus of variations allows me to characterize socially efficient adjustment path and compare it with the decentralized outcome.

It is hoped that methods developed in this paper may be found useful in subsequent research on short run dynamics, not necessarily in area of transition.

My economic universe rests on two assumptions: imperfect labor markets, and trade-off between special and general human capital due to differential complementarities across skill types. The extent of specialization in planned economies of Central and Eastern Europe prior to transition, undeveloped labor market infrastructure, and the size of the shock that transformed the structure of relative rewards, make those countries an ideal testing ground for my theory of adjustment. Meta-evidence on wages and unemployment incidence of workers with general and vocational secondary education appears to be consistent with my model. East Germany makes for a polar case of a totally different evolution of wage differences. I show how German unification of 1991 may have altered trade-offs faced by workers compared to other CEE countries, and led to much higher transitory output decline and larger unemployment (which are hard to account with existing theories of transition).

My model may also be helpful for understanding the role of flexibility and specialization in generic changing environments, with lessons for growth and development economics.

Genealogically, my paper is related to three areas of research: equilibria in economies with heterogeneous workers, complementarities across workers, and implications for sorting and inequality [Shimer and Smith, 2000; Burdett and Coles, 1997; Lu and McAfee, 1996; Kremer and Maskin, 1996; Kremer, 1993], research on differential productivities and endogenous job creation in search framework [Acemoglu, 1996a; Saint-Paul, 1996], and analysis of reallocation dynamics [Aghion and Blanchard, 1994; Caballero and Hammour, 1996; Castanheira and Roland, 1997; Brixiova and Li, 1998]. Shimer and Smith [2000], Burdett and Coles [1997], and Lu and McAfee [1996] exemplify a growing body of literature analyzing steady states in a rich class of search models with heterogeneous agents, with implications for sorting. Kremer [1993] and Kremer and Maskin [1996] address similar questions in static Walrasian framework, where production technology is characterized by complementarities and differential sensitivities of output to workers with different productivities. Kremer and Maskin [1996] make very strong assumptions on functional forms to obtain patterns of sorting by skill and evolution of across-group inequality similar to those observed in industrial countries. My paper, in contrast, focuses on out-of-steady-state dynamics; it only uses a minimal set of technological assumptions; most of the results are driven by the dynamic structure of the model and the frictional nature of the labor market. Acemoglu [1996a] and Saint-Paul [1996] analyze endogenous structure of jobs in search framework when labor force is heterogeneous. Their analysis is also confined to comparative statics of steady states. Aghion and Blanchard [1994] and Caballero and Hammour [1996] analyze the reallocation process in transition. They rely on either ad hoc assumptions or numerical analysis to characterize dynamics of adjustment and efficiency of decentralized path. My approach is analytical, even though I study a higher-dimensional labor market.

Formal analysis of this paper is motivated by factual evidence on skill obsolescence, excessive specialization, and incompatibilities in transition economies [Boeri, Burda and Köllõ, 1998; Bordin, 1998; Cazes and Scarpetta, 1998; Flanagan, 1998; Rutkowski, 1996; Rutkowski, 1998a; Rutkowski, 1998b; Orazem and Vodopivec, 1997; The World Bank Development Report, 1996]. Orazem and Vodopivec [1997], Flanagan [1998], and Rutkowski [1998a], among others, find that returns to human capital rose rapidly in most countries of the Central and Eastern Europe (CEE). At the same time experience profiles have flattened significantly. Skill premia were not shared by all age groups. Boeri et al. [1998], Cazes and Scarpetta [1998], Rutkowski [1998b], and Orazem and Vodopivec [1997] document that vocational degrees are associated with higher unemployment incidence and longer unemployment durations. Bordin [1998] provides anecdotal evidence on incompatibilities among workers with different backgrounds, in particular, in foreign-managed firms.

However, none of the existing empirical studies explore systematically the differences and changes in labor market prospects of workers with general and special skills. My model provides a framework to analyze these questions in a structured way.

Aggregate evidence suggests that East Germany has a very different pattern of transition than the other CEE countries. Higher transitory unemployment and larger output movements in this country seem to be consistent with the predictions of my model. I estimate returns to different types of schooling for young German workers and indeed find significant growth in vocational premium over 1992-1997. The contrast between East Germany and other countries with socialist legacy is indicative of the dynamic effects identified in this model.

In what follows, I briefly explicate the distinction between special and general human capital in section 2. Section 3 presents a very simple model that incorporates and highlights some of the central features of my analysis. The fully fledged model and results are presented in section 4. Section 5 discusses some facts about labor markets in transition economies which seem to be consistent with my theory. In particular, East German experience appears to be an example of swift transition driven by early specialization. Section 6 concludes by discussing policy implications and avenues for future research. Appendices contain detailed formulation of the analytical framework and proofs of the main results.

2 Specialists and generalists

Usually in economics, human capital is measured by a single real number designed to grasp the *expertise* of a worker in some special skills. One can think of at least one more dimension – *flexibility* of the human capital, or the broad ability of a worker or professional to perform *multiple* different tasks, or work in a number of different organizational environments. There is a non-trivial trade-off faced by all those who retrain in the middle of their career as well as by new entrants into the labor force between the degree of specialization and its extent: narrow specialization substantially raises the marginal productivity of worker in a specific task, while broad specialization raises marginal productivity in many different tasks, though not on a par with the expertise of a narrow specialist. In any single task dab hands outperform Jacks of all trades, but in a frictional labor market Jacks have the advantage of flexibility. The only model I am aware of that describes workers confronted with the choice of general vs. specific training was presented in [Kim, 1989], but it limits discussion to a static partial equilibrium context with no interaction of workers.

There are, in fact, two reasons why a specialization in a narrow task may be undesirable. There may be perfect labor markets, but an *aggregate* uncertainty about which single specialization will be profitable in the future. Since investment in human capital is irreversible, risk-averse agents making decisions before the resolution of uncertainty will prefer to invest slightly in a number of skills, rather than one. Another reason why specialization may be undesirable stresses the *individual-specific* uncertainty about a match under labor market imperfections. This is the story I want to tell.

In the absence of aggregate uncertainty, the more perfect are the labor markets, the smaller is the flexibility advantage. In frictionless markets workers match perfectly with other workers and factors. It is in the imperfect market environments that dab hands have disadvantages. Thus there are strong grounds to believe that in transition economies, where limited labor mobility, high degree of segmentation, and undeveloped infrastructure are central features, the specialization trade-off should assume primary importance.

3 A simple myopic model of human capital adjustment

Before presenting a full-fledged RE model, in order to build intuition and introduce the relevant trade-off, I present an extremely simplistic myopic model of aggregate human capital adjustment. Section 4 will soon correct for all deviations from rationality and realism this section makes.

There are three potential types of workers, $i = 1, 2, 3$. Initially the entire population consists of type 1 workers. Type 1 are workers highly skilled in obsolete technique[1]. Type 3 are those highly specialized in new technique. Type 2 are the Jacks of all trades. Workers can only produce output in pairs. The output of a pairing of workers i and j is given by $y_{ij.}$

I make the following assumptions about relative productivities of various matches:

monotonicity[2]: $\quad y_{13} < y_{11} < y_{12} < y_{22} < y_{23} < y_{33},$
and
increasing returns to specialization $\quad y_{33} - y_{23} > y_{23} - y_{22}$

Workers are risk-neutral and infinitely-lived. Time is discrete. A proportion ρ of all workers, drawn randomly and uniformly in each period, get a chance to retrain, i.e. change their type arbitrarily. The proportions of types in labor force are denoted by p_j.

Workers find a partner at the beginning of a period, produce output, and part at the end, without taking each other's email address. Therefore, they engage in search of a new

[1] I use the word "technique" instead of "technology" to stress that not only technological shock can make skills obsolete. New forms of contracts becoming available or changes in relative prices of goods are other examples of such a shock. The word "technique" in this context simply refers to the best available way to produce value.

[2] The results still hold under weaker assumption $y_{11} < y_{13} < y_{12}$.

partner each period. I assume that the search is infinitely efficient in the sense that everyone gets a partner, whose identity, however, is randomly drawn from the pool of all workers. In this sense, matching is random.

I denote by w_{ij} the wage paid to i in a match (i, j). In what follows I will assume for simplicity that a pair of workers simply split their output evenly:

$$w_{ij} = w_{ji} = y_{ij}/2.$$

To put the implications of the differential complementarity assumptions in stark light, I choose the simplest way possible to describe the intertemporal decision-making. I assume that workers make choices based entirely on contemporaneous costs and benefits.

In each period workers who get retraining opportunity, compare the expected wages they would realize with each set of skills. Type i expects to earn the amount $\bar{w}_i = \sum_{j=1}^{3} p_j w_{ij}$. It is obvious that no one would ever choose to obtain type 1 training, because it is strictly dominated by upgrading to type 2:

$$\sum_{j=1}^{3} p_j \left(w_{2j} - w_{1j}\right) > 0 \text{ because } w_{2j} > w_{1j} \forall j.$$

What about getting the new specialized skills (type 3)? Initially, $p_1(0) = 1$, $\bar{w}_3|_{t=0} = w_{31} < w_{21} = \bar{w}_2|_{t=0}$. So, the workers will not choose type 3. The proportions are changing according to difference equations: $p_1(t) = (1-\rho) p_1(t-1)$, $p_2(t) = p_2(t-1) + \rho p_1(t-1)$.

What happens when type 2 workers populate the market, $p_2 \approx 1$? $\bar{w}_3 = w_{32} > w_{22} = \bar{w}_2$. So, when market gets dominated by generalists, type 3 becomes the type of choice. By continuity, there exists time s_0 when it first becomes attractive to specialize:

$$\begin{aligned} \bar{w}_3(s_0) = p_1(s_0) w_{31} + p_2(s_0) w_{32} &\geq p_1(s_0) w_{21} + p_2(s_0) w_{22} = \bar{w}_2(s_0), \\ \bar{w}_3(s_0 - 1) &< \bar{w}_2(s_0 - 1). \end{aligned}$$

There are following lessons from the simple model:

1. Along the path to the long run steady state, the economy first accumulates and then decumulates a buffer stock of workers with "bridging" human capital, which is compatible with both old and new specialized skills.
2. If the investment in general skills is impossible, the economy may never embark on adjustment.
3. There is an inverse U-shaped path of income inequality in transition; inequality temporarily rises both across and within groups.

4 A rational expectations model of human capital adjustment

4.1 Assumptions of the model

Here I add more realism to the simple model of the preceding section by allowing for unemployment and job search, as well as endogenous matching decisions. It is a variation on the job-search models. The classic treatment of these models is contained in [Pissarides, 1990]. However, in order to make model tractable, I need the same major simplification as in section 3: abandonment of the firms' side altogether[3]. There will be neither firms, nor vacancies in the model. The workers match among themselves directly.

There are two alternative interpretations of the model with small variations in the assumptions. The main interpretation holds that workers are infinitely-lived. From time to time, at rate ρ, they get a chance to update their skills. An alternative interpretation is that workers retire at rate ρ, and are immediately replaced in the labor force by their offspring. Which version is adopted makes no difference for formal results, however, the interpretation of results and their implications differ accordingly. I call the former the "infinite life" interpretation, and the latter the "retirement interpretation".

4.1.1 Payoffs

The model is in continuous time. The production side (outputs and wages in pairs (i, j)) is the same as in section 3, but I now allow for a possibility of unemployment. Let z denote the flow return to unemployed agents. At this point, I do not specify where it comes from (utility of leisure, unemployment benefits, home production). This will become more important when discussing efficiency issues. I assume that z is a constant.

The wage-setting is as in section 3:

Assumption 1.

$$w_{ij} = w_{ji} = y_{ij}/2, \tag{1}$$

This specification may look too rigid, but there may be many potential reasons why wages are inflexible, for example, sociological reasons [Bewley, 1995; Bewley, 1998], or efficiency wages considerations. I chose this assumption, rather than Nash bargaining solution, often used in job search models, because it ensures that results are not driven by peculiarities of the latter. I make remarks about implications of alternative wage setting specifications later in the paper. In fact, any wage-setting procedure that does not completely insulate some workers from the adversities of random pairing, would give rise to qualitatively similar results. What is also important, assumption 1 makes wage vector constant, which considerably simplifies the analysis.

Workers' utility is linear in consumption.

4.1.2 Flows in and out of unemployment

At a moment in time, a worker i can be working in a pair with type j, or be unemployed and searching for a match. I denote the unemployment state $(i, 0)$, so there are now four possible employment states $(i, 1), (i, 2), (i, 3), (i, 0)$.

There are following flows between the employment states. When an agent is unemployed, in every period a *potential* match with another unemployed worker arrives at rate[4] m. Equiv-

[3] I worked out description of model dynamics with firms in [Sarychev, 1999].

[4] Note that in our context m is a matching constant analogous to the matching *function* of job-vacancy models. This happens because workers match among themselves; the proportion of pairs created is a general function $m(U)$. Assuming that this relationship is homothetic of degree 1, we obtain $m(U) = Um(1) = Um$.

alently, this means that waiting time before finding a match in the unemployment pool is distributed exponentially with parameter m. The *type* of the match is also random with probability of matching with type j equal to their proportion in the entire unemployment pool:

Assumption 2. *Matching in the labor market is completely random.*

I make remarks about relaxation of this assumption in subsection 4.7.3. An additional motivation for this assumption in the "retirement" interpretation is that there is imperfect substitutability between young workers and old workers as in [Kremer and Thomson, 1998], so agents cannot sort by age.

Denoting $U_j(t)$ the number of unemployed type j workers at time t, we compute the total unemployment as $U(t) = \sum_{j=1}^{3} U_j(t)$ and define the proportions $q_j(t) = U_j/U$. In what follows I will omit the time index for most time-varying variables when it is safe to do so. Summarizing the discussion, the matching rates are therefore mq_1, mq_2, mq_3. If no match has arrived, worker remains in the unemployment pool.

Working relationships dissolve for exogenous reasons at a rate δ, upon which workers find themselves in the unemployment pool.

4.1.3 Choice variables

So far what happens is completely out of agents' control. Let us now turn to training and retirement. As mentioned earlier, there are two interpretations of the model. In the "infinite life" interpretation, once in a while when unemployed, at rate ρ, worker gets an opportunity to retrain, that is to change own type arbitrarily (or choose to remain the same type). The model may be equivalently reformulated to have dynasties of agents. In this "retirement" interpretation, type of agent is chosen once and for all at the beginning of her/his career; ρ is the retirement rate. Upon retirement of an adult worker, offspring of the retired person enters the labor market. To avoid excessive stress on *re* in retraining, I will use the less specific term *training* below. Training is instantaneous and costless[5]; following it, the worker transits to the respective unemployment pool.

In contrast to the model of section 3, where workers always had *some* match and future was irrelevant for wage decisions, here those who for any reason break with their partner will anticipate a positive waiting time until they find another match. In a market with frictions, there is always a positive option value of preserving a match. Even though assumption (1) makes compensation of all workers independent of their reservation utility, it is not in violation of individual rationality constraints. The model allows for voluntary dissolution of a match if one worker's value of staying falls below her/his reservation utility (expected utility when unemployed). Agents thus have an additional set of choice variables ϕ_{ij}. ϕ_{ij} is a binary variable indicating a positive decision of type i to work with type j. If $\phi_{ij} = 1$, worker of type i agrees to stay in productive partnership with type j worker. If $\phi_{ij} = 0$, we say that type i decides not to match with j (decreasing, obviously, one's probability of finding a match). I assume that when a type i worker changes one's ϕ_{ij} while employed with type j, she/he transits into the unemployment state instantly at will.

[5] I relax the assumption of free training in section 4.7.1.

4.1.4 Other notation and Bellman equations

The additional notation is as follows: $V_{ij}(s)$ is the expected discounted utility of type i working with type j (or unemployed if $j = 0$) at time s.

In the search literature Bellman's dynamic programming approach plays a prominent role. While its usefulness is somewhat limited for my analysis, I formulate the Bellman asset value equations to facilitate comparisons with that literature. I present these equations for the "infinite life" interpretation:

$$rV_{ij} = w_{ij} + \delta\left(V_{i0} - V_{ij}\right) + \dot{V}_{ij}, \forall i, j > 0 \tag{2}$$

$$rV_{i0} = z + \rho\left(\max_k \{V_{k0}\} - V_{i0}\right) + \sum_j mq_j\phi_{ij}\phi_{ji}\left(V_{ij} - V_{i0}\right) + \dot{V}_{i0}, \forall i \tag{3}$$

The flow equations in general form are

$$\dot{U}_i = \frac{\delta}{2}L_i - mU_i\sum_j q_j\phi_{ij}\phi_{ji} - \rho U_i + \mathbb{I}\left(j = \arg\max_k \{V_{k0}\}\right)\sum_{k=1}^{3}\rho U_k \tag{4}$$

$$\dot{L}_i = -\frac{\delta}{2}L_i + mU_i\sum_j q_j\phi_{ij}\phi_{ji} \tag{5}$$

In fact, the "retirement" interpretation is more expensive in terms of notation, because it must keep track of four more state variables (number of all pairs (i, j) among the employed). I do not write the Bellman equations for this interpretation here. Despite little variations in assumptions, formal results hold under either interpretation of the model.

The construction of the model in terms of transition probabilities is developed in appendix A.

4.2 Informal description of the adjustment path

Most contributions to search literature are limited to the analysis of steady states. In contrast, my focus is on the adjustment dynamics. Characterizing the steady state in my model is trivial. In the long run all the workers will be of type 3. The unemployment rate is computed in a straightforward fashion from the condition on stability of flows into and from the unemployment. It is how the economy gets to that steady state that interests me.

Before proceeding with formal analysis, I describe informally the equilibrium dynamics. I am interested in a rational expectations path. That means that all agents correctly anticipate the evolution of economy aggregates. In fact, while the individual transitions between employment states are random, the aggregate dynamics are deterministic if all agents play deterministic strategies. At the beginning of the time all agents are of type 1. Initially, there is full employment [6]. There comes a sudden shock that makes available a new production technique. For every agent faced with a decision to retrain there is no doubt that choosing

[6] This assumption is more in line with the transition story, where the full employment was politically enforced. Letting the unemployment at time 0 to hover about some natural rate does not change the qualitative predictions.

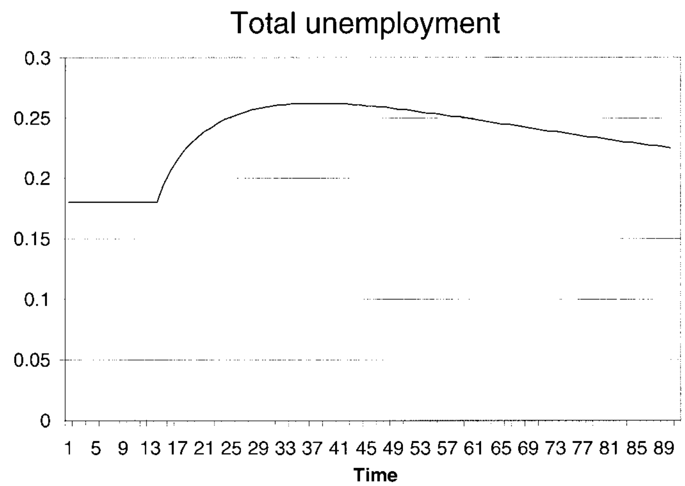

Figure 1:

type 2 is better than remaining type 1 (type 3 might be even more attractive for some parameter values, but I rule out this possibility until later). This is so because type 2 workers are rewarded with $w_{21} > w_{11}$ when matched with a type 1 worker, and with $w_{22} > w_{12}$ with a type 2 worker, while probabilities of finding either type of match are the same for type 1 and type 2 worker. Thus in the short run being type 2 worker is strictly superior (this will only strengthen with time when the possibility of labor market isolation for type 1's arises). Therefore, everyone who gets a chance, transits through training into the unemployment pool of type 2 workers. Why choosing type 3 might initially be a bad idea? Because (similarly to the model in section 3) in pairs with obsolete specialists, the modern specialists (type 3) are making only a meager $w_{31} = y_{13}/2 < y_{11}/2 = w_{11}$. Initially, the labor market is dominated by the obsolete specialists, and therefore earning prospects of type 3's are mediocre. They will improve, however, as the presence of type 2 workers in the market is increasing. Suppose an agent who goes through training at time s_0, decides to become type 3. At all later times, by monotonicity, all agents in training must choose type 3 also, because if anything, the labor market prospects for type 3 have improved compared to time s_0. Thus there will be a distinct threshold moment on the equilibrium path, at which the fortunes of type 2 and 3 workers flip. Only after that moment the transition truly gains momentum.

Things do not go smoothly on the road to prosperity. As their market presence grows, generalists may find that working with type 1s is no longer preferable to being unemployed. The opportunity cost of working in a pair with type 1 is the forgone prospect of making a better match, the prospect that grows over time. At some point type 1 workers can only match within their own constantly shrinking pool.

Type 3s may later become similarly reluctant themselves to match with type 2 workers.

Along the adjustment path, there is an abnormally high unemployment (see Figure 1), due to the fact that some types do not match. By avoiding a subset of prospective coworkers, those types increase their average waiting time.

Moreover, there is a rise in long-term unemployment, when obsolete specialists become

Figure 2:

isolated (see Figure 2). Due to the very small effective pool of prospective partners, their outflow rate to jobs is much lower than that of workers with more up-to-date skills. Long-term unemployment goes down gradually as human capital is replaced .

In the rest of this section I will study the dynamics of adjustment to the long-run steady state. To do that I analyze the optimal decision thresholds, or switching times.

4.3 Equilibrium thresholds

I analyze my model in terms of *thresholds*. In the adjustment context, I use the term "threshold" to designate moments in time when agents change their binary strategies regarding training and matching. Why are there thresholds? Because the relative attractiveness of one choice over another is locally monotone in my model. If over a period of time $[s, t]$ workers have been (optimally) training as type 3, then from moment t onwards all workers must also choose type 3, since the skills distribution has further improved during that period.

The most important threshold is the moment when all agents in training start choosing type 3: s_0. I alternate between calling s_0 the specialization threshold and the moment of specialization takeoff.

Apart from the choice of type, there are also thresholds in the matching choice ϕ_{ij}.

Workers of type 2 will prefer to forego a match with type 1 if being unemployed and searching gives them higher expected utility. s_{obs} is the earliest time generalists decide not to match with type 1 (I call this the *obsolescence* threshold), s_{negobs} is the latest time they avoid matches with type 1s. These thresholds may well be zero or $+\infty$. Notice that $s_{negobs} \leq +\infty$ in general. It may happen that after a spell of discriminating against type 1, workers of type 2 decide to match with them again, if their own condition sufficiently deteriorates.

s_{seg} is the earliest time type 3 workers decide not to match with type 2 (I call this the *segregation* threshold).

When workers of type 1 are avoided by type 2, I say that the former are *isolated* in the labor market, because they are left to match only among themselves. When workers of type 3 pass type 2, I say that the former are *segregated* in the labor market.

In what follows I will assume that type 3's *never* match with type 1's, to sharpen the focus on their incompatibilities, although working out the generalization with endogenous ψ_{13} is straightforward[7].One might also argue that potentially there are other separation thresholds. Namely, type 2 worker may decide not to match with any type in order to hasten the arrival of a training opportunity. I do not study these other thresholds for economy considerations, as doing this using the methods developed here is straightforward.

4.3.1 Conditions on thresholds

Consider a decision of a worker who at time s just received a training opportunity. Upon completing training, agent transits into unemployment pool of either type 2 or 3. The relevant choice is between type 2 and type 3 unemployment states. To be at a specialization threshold, she or he must be indifferent between continuing her (his) career as either type. The variables that summarize the expected discounted utility under all foreseeable contingencies in the two states are V_{20} and V_{30}. Consequently, at the threshold

$$V_{20}\left(s_0\right) = V_{30}\left(s_0\right). \tag{6}$$

I call the time path of $(V_{20} - V_{30})$ the relative attractiveness schedule.

Workers of type 2 will prefer to forego a match with type 1 if being unemployed and searching gives them higher expected utility: $V_{21} < V_{20}$. Thus, the moment s_{obs} satisfies

$$V_{21}\left(s_{obs}\right) - V_{20}\left(s_{obs}\right) = 0 \tag{7}$$

Workers of type 2 may later revert to matching with type 1, and the moment they do so, s_{negobs}, satisfies the same condition (7). The difference between the two is in the way $(V_{21} - V_{20})$ crosses the zero line: from above or from below (see Fig. 3). In principle there might be more than 2 obsolescence thresholds if the schedule $(V_{21} - V_{20})$ crosses zero more than twice. I have no predictions about the exact number of thresholds. They all, however, behave in the fashion, described by propositions of section 4.5.

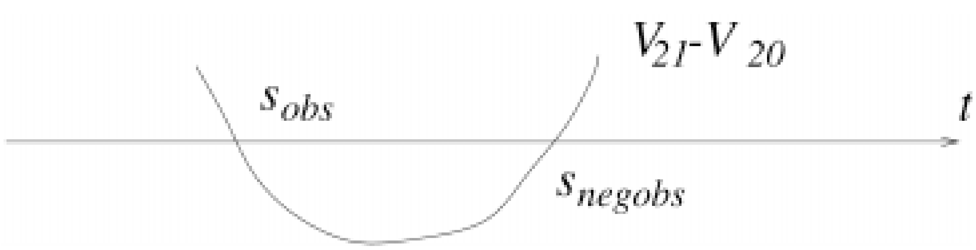

Figure 3:

[7]Type 3s will never match with type 1s endogenously if $y_{13} \leq 2z$, and in fact, if y_{13} is higher than $2z$ but not by much.

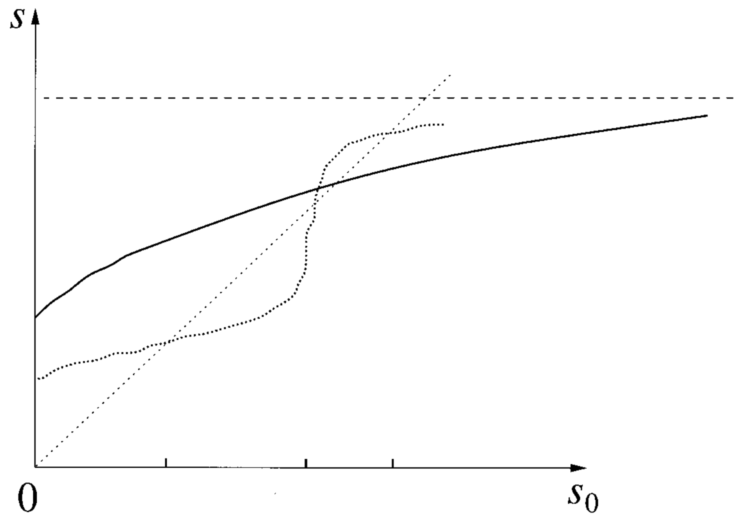

Figure 4: Multiple thresholds

Moment s_{seg} is when the new specialists (type 3) start avoiding matches with type 2. It is given by the equation $V_{32} = V_{30}$, or

$$V_{21}(s_{seg}) - V_{20}(s_{seg}) = 0 \tag{8}$$

4.3.2 Relative timing of thresholds

While in all meaningful cases $s_{seg} \geq s_0$, the model does not specify, whether $s_{obs} < s_0$ or $s_0 < s_{obs}$. In fact, both situations can be potentially observed on an equilibrium path, but I have nothing to say about the determinants of relative timing of thresholds. The results below are unaffected by this ambiguity.

4.4 Uniqueness of equilibrium threshold

Before presenting my comparative dynamics findings, I state the result on the uniqueness of the specialization threshold.

Proposition 1 *Equilibrium threshold s_0 is uniquely defined by condition (6).*

Proof. Notice that equilibrium threshold is a fixed point of the individual best response correspondence $s(s_0)$, where s is the individual threshold. A sufficient condition for equilibrium uniqueness is that for any t, $t' > t$, $s(t') - s(t) < (t' - t)$ (see Fig. 4)

Consider the relative attractiveness of specialization at time t if the individual threshold is s and aggregate threshold is s_0: $M(t, s, s_0) = V_{20}(t, s, s_0) - V_{30}(t, s, s_0)$. I will assume that M is continuously differentiable for simplicity, although the result holds for non-differentiable functions as well. Since s solves $M(s, s, s_0) = 0$, $\frac{ds}{ds_0} = -\frac{\partial M/\partial s_0}{\partial M/\partial s}$. To prove the uniqueness of

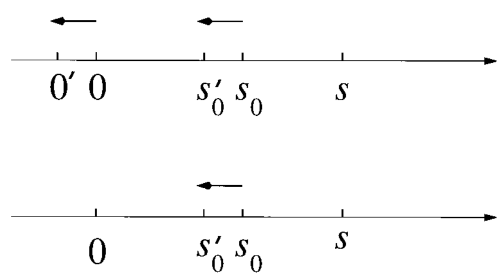

Figure 5: Change in individual threshold vs. change in aggregate threshold.

equilibrium, it suffices to show that $\frac{\partial M}{\partial s_0} < \left|\frac{\partial M}{\partial s}\right|$. Consider small $h > 0$ and contrast changes in M:

$$M(s, s_0 - h) - M(s, s_0)$$
vs.
$$M(s + h, s_0) - M(s, s_0)$$

I argue that $M(s + h, s_0) - M(s, s_0) > M(s, s_0 - h) - M(s, s_0)$. To see this, freeze t (similarly to tying coordinate system to a moving object in high school physics), and consider what happens when s increases by h, or s_0 decreases by h (see Fig. 5). In the former case two important dates shift to the left – the time when type 2s start pouring in, and the time when type 3s appear. Both of these shifts increase the relative attractiveness of specialization M. In contrast, in the latter case, only one of the two dates shifts, and the other stays. Obviously, the effect on M is larger in the case when s increases by h.

Since the reasoning is valid for any t and any h,

$$-\frac{\partial M}{\partial s} = \left|\frac{\partial M}{\partial s}\right| > \frac{\partial M}{\partial s_0}.$$

□

At the heart of the uniqueness result lies the greater sensitivity of relative value of specialization to variation in own threshold vs. variation in collective threshold. Suppose we have multiple equilibrium trajectories, with thresholds s_0 and s_0', $s_0' > s_0$. Moving both thresholds from s_0 to s_0', the increase in agent's relative value of specialization due to postponing individual threshold exceeds the decrease due to the postponement of the aggregate threshold. Therefore 6 cannot hold at both dates s_0 and s_0'.

This result on the uniqueness crucially relies on the assumption of costless training. When it is relaxed, there are potentially multiple equilibria, in one of which there may not be a specialization threshold at all. For details, see subsection 4.7.1.

4.5 Comparative dynamics

In this section I study the behavior of optimal thresholds when some of the parameters vary. This is an exercise in comparative dynamics, because I consider different adjustment paths. not steady states.

Appendix A develops an analytical approach toward optimal thresholds. Explicit handling of the Markov transition matrix and the associated generator matrix allows expressing state-specific utilities $V_{ij}(t)$ in a non-recursive way. As a result, V_{ij} can be differentiated to yield predictions about comparative dynamics. Supermodularity theorem 1 ensures that the results hold unconditionally in general equilibrium.

In the propositions below, the relevant experiment is to consider two almost identical economies which differ in only one parameter of interest. The comparison then is between two equilibrium paths. A convention: I say there is a negative effect on a threshold, if it shifts to the left on the time axis, and vice versa for a positive effect. All proofs are relegated to appendix.

Let us first juxtapose economies with different wage vectors.

Proposition 2 *An increase in w_{21} leads to postponed specialization takeoff (higher s_0). Isolation of type 1 workers occurs later along the path and lasts shorter (higher s_{obs}, lower s_{negobs}). Although there is no* direct *effect of w_{21} on the segregation threshold, the general equilibrium effect is positive, too.*

An increase in w_{21} leads to a shift of the relative attractiveness schedule $(V_{20} - V_{30})$ upward, at any given q_2 reducing the incentive to specialize. On the other hand, a reduction in w_{21} has an opposite effect, so the specialization takes off earlier. The intuition behind the result is that with higher w_{21}, type 2 workers earn more in expectation, whereas type 3 workers' expected wage is unaffected.

Now consider the effects on the obsolescence thresholds. An increase in w_{21} shifts the $(V_{21} - V_{20})$ schedule upward. The obsolescence threshold s_{obs} shifts to the right, because $(V_{21} - V_{20})$ is decreasing in its neighborhood; meanwhile s_{negobs} shifts to the left, as $(V_{21} - V_{20})$ is increasing at that point (see Fig. 6). Both shifts imply a shorter isolation spell of type 1 workers.

The result is intuitively appealing: a lower wage in the match with obsolete workers makes discrimination in matching pay off. An increase in w_{21} leads to a longer spell of type 1 and type 2 agents working together.

A real-world comparison corresponding to proposition 2 is between two economies, one of which exhibits excessive specialization. An economy, whose stock of human capital is incompatible with market realities, has little to hold on to. It therefore adjusts more aggressively than the economy in which generalists have some flexibility advantage. Obviously, this does not mean that welfare is higher on the transition path for the more extinct economy. As differentiation of the expression for welfare in appendix B would show, a decrease in w_{21} makes the society worse off.

Proposition 3 *An increase in unemployment benefit z leads to earlier specialization, earlier isolation of type 1, and earlier segregation of type 3 workers (s_0, s_{obs}, s_{seg} all go down).*

This accords with intuition: higher benefits make job searchers more discriminating and cushion the adverse effects of labor force composition on potential type 3 specialists. Thus, unemployment benefits may serve as an instrument in affecting the *speed* of transition, not only the welfare of the unemployed, even with fixed search intensities.

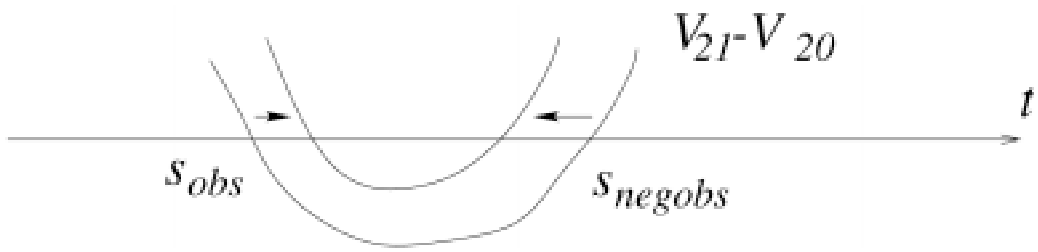

Figure 6:

Interpreting z as home production, this proposition implies that primitive economies adjust faster to technological or organizational shocks than the economies with greater importance of the formal sector. A comparison between growth rates in China and Central Europe may be a case in point (but China also has a much larger ratio of y_{33}/y_{11},the distance to steady state).

Consider now how the threshold in (19) changes when parameters ρ, δ, and m vary.

Proposition 4 *An increase in training opportunity arrival rate (retirement rate) ρ has an ambiguous effect on the specialization threshold. The incentive effect is positive, while the speed effect is negative. Asymptotically, the negative speed effect dominates.*

The total impact of a given change in ρ can be decomposed into incentive effect and speed effect. The former is the effect of the change in the *individual* transition rates with aggregate dynamics unchanged (higher ρ implies shorter expected life span, or in infinite life interpretation, better odds of getting training while unemployed). The latter is the impact of the altered evolution of the aggregates.

In the context of proposition 4, the incentive effect is positive. In the "infinite life" interpretation, higher odds of getting a training opportunity in the future *increase* the option value of conservative policy (choosing type 2 instead of type 3). Improved access to education leads to slower adjustment! It is interesting to notice how the same effect, depending on the interpretation of the model, may seem either entirely conventional or somewhat unexpected. In the retirement "interpretation" of the model, higher retirement rate ρ implies a larger effective discount rate. As a result, agents care less about distant future, choosing the immediately rewarding type 2.

The speed effect is negative, because with higher rate of training (generational turnover in the "retirement interpretation"), there will be more quality partners in the future, making type 3 more attractive.

The total effect is likely to be negative if ρ is large enough already, discount rate r is high, or matching rate m is high. Notice, also that higher ρ unambiguously speeds up the initial accumulation of type 2s. As ρ grows asymptotically, training opportunity arrives instantly and all workers have an incentive to invests in special skills at time 0. Thus, asymptotically, speed effect dominates the incentive effect.

Proposition 5 *An increase in separation rate δ leads to postponed specialization threshold: s_0 goes up.*

The incentive effect of a change in δ is zero: both types expect to stay unemployed proportionately longer. The speed effect is positive. Due to the higher rate of match dissolution, there will be a relative increase in the proportions of type 1 and 2 workers on the market because they account for a larger share of employment. This makes specialization less attractive. Figures 7 and 8 depict trajectories of unemployment and output per capita for high and low δ (dashed line corresponds to a higher separation rate).

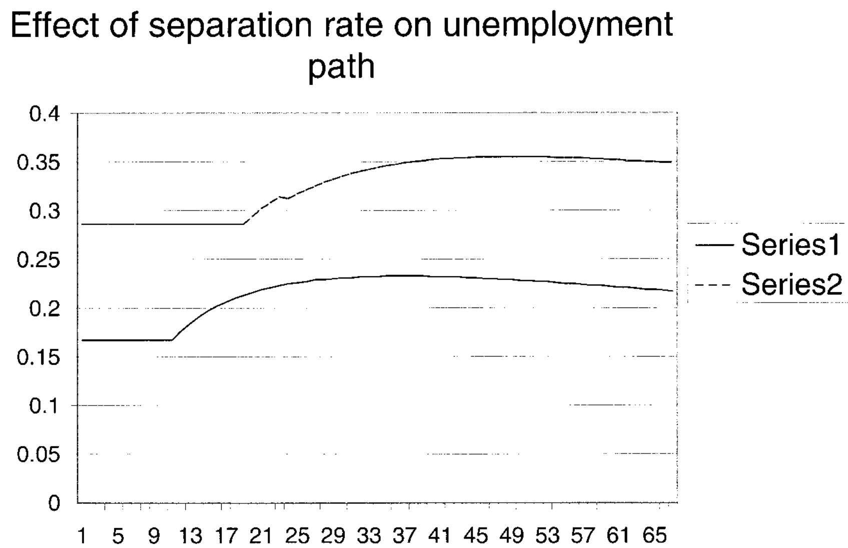

Figure 7:

Economies with exogenously higher turnover (lower tenures) will experience slower adjustments to shocks, with higher transitory unemployment, according to the model.

Proposition 6 *The incentive effect of an increase in matching rate* m *is negative* (s_0 *goes down). The speed effects is also negative if at least half of the unemployed population is of type 1 at the specialization threshold.*

An increase in m acts through expectation of shorter searching time in the future (incentive effect) and increased proportion of suitable partners (speed effect). The latter effect follows because with higher m, unsuitable partners (type 1) match faster when they still dominate the market, and their proportion in unemployment falls. Asymptotically, the incentive effect dominates: as the labor market becomes progressively more efficient, there is less dynamic inertia associated with adjustment.

4.6 Efficiency

Let us now turn to the questions of constrained efficiency of the transition process. Suppose there were a social planner who would be able to dictate agents what kind of training to undertake, and oversee the separation decisions in individual workplace. Equivalently put,

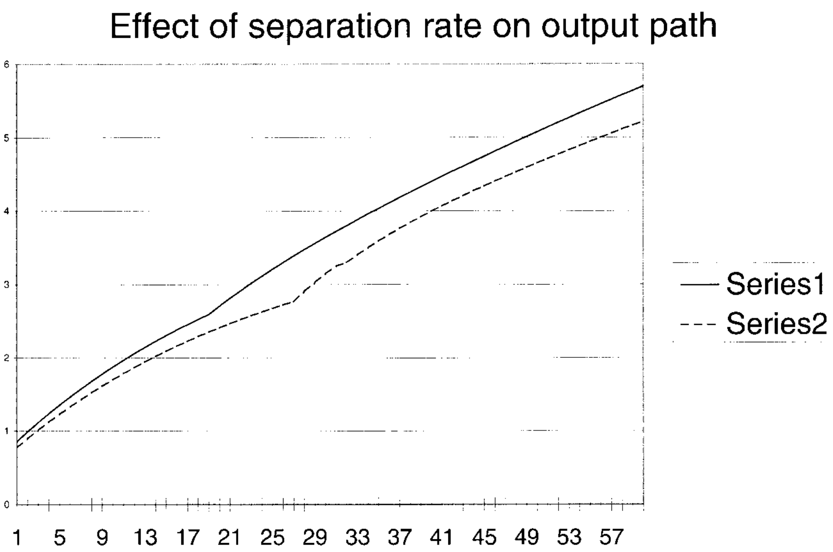

Figure 8:

the social planner can choose the specialization threshold as well as other thresholds. The solution to planner's problem is a constrained Pareto optimum subject to labor market frictions. The optimal command threshold is computed in appendix B under the utilitarian welfare criterion.

Appendix B demonstrates that individual decisions to specialize create an externality on training choices of other agents. The earlier a subset of agents decide to specialize, the more attractive specialization choice becomes for everybody. Hence, in decentralized equilibrium, specialization takes off too late due to the coordination failure among workers.

An additional source of inefficiency of decentralized equilibrium is misallocation of production. Earlier specialization means there are more type 3 workers on the market which in turn imply more matches of type $(2,3)$ and fewer matches of type $(1,2)$ and $(2,2)$. Higher output in pairs with modern specialists compensates for the output loss in obsolete and semi-obsolete matches, so the society as a whole benefits. On the distribution side, however, earlier specialization hurts workers with obsolete skills by prolonging their waiting time, and benefits workers with general skills.

The inefficiency of the decentralized equilibrium path makes a case for government intervention. Any policy that can speed up specialization will be locally beneficial. Section 4.5 seemingly has a few other, counterintuitive implications for welfare-enhancing policies. I discuss them in the section 4.8.

4.7 Extensions

4.7.1 Cost of education

Until now I assumed that training was costless for the agents. Relaxing this assumption does not lead to significant revision of the results. If C_2 and C_3 denote the costs of training as type 2 and 3 respectively, there may be now two specialization thresholds instead of one, depending on what interpretation of training we assume. In the "infinite life" interpretation, agents retrain from time to time. Someone who has already acquired type 2 does not have to pay the training cost C_2 again if she does not want to change her type, but agent of type 1 must pay this cost. Consequently, for agents of type 1 the specialization threshold will assume the form

$$V_{20}(s_0) - V_{30}(s_0) - C_2 + C_3 = 0, \tag{9}$$

while for type 2 agents it will be

$$V_{20}(s_0) - V_{30}(s_0) + C_3 = 0. \tag{10}$$

In the "retirement" interpretation of the model workers acquire their type only once, at the beginning of their career. Hence, the threshold is unique and given by (9).

It is immediate that the propositions 2–6 are unchanged. An additional straightforward result is

Proposition 7 *An increase in C_3, or conversely, decrease in C_2, pushes up all specialization thresholds.*

Proof. Differentiate (9) w.r.t. $C_3 : \frac{d}{dC_3}(V_{20} - V_{30} - C_2 + C_3) = 1 > 0$. Thus the relative attractiveness schedule shifts up. Similarly for C_2. □

If specialization is costly relative to general training, it will be postponed. At the same time unemployment will be lower on the adjustment path (due to matching effects).

An additional effect of introducing costly education into the base model is a potential for multiple equilibria. Consider a situation when labor force consists of type 2 workers only. In absence of costs of training, every agent undergoing training would choose to specialize because the discounted utility of being type 3 and unemployed, $V_{30} = V_{30}^{high} > V_{20}^{high}$. V_{30}^{high} is the value this utility takes under the expectation that all workers will specialize. In contrast, if no one specializes, utility in state $(3,0)$ is $V_{30} = V_{30}^{low}$ which is still higher than V_{20}^{low} because $w_{32} > w_{22}$. However, $V_{30}^{low} - V_{20}^{low} < V_{30}^{high} - V_{20}^{high}$ due to complementarities ($w_{33} > w_{32},$). Now introduce the costs of training in the amount C_3, $V_{30}^{low} - V_{20}^{low} < C_3 < V_{30}^{high} - V_{20}^{high}$. There will be an equilibrium adjustment path on which everyone decides to specialize instantly (because $V_{30}^{high} - V_{20}^{high} > C_3$) and an adjustment path on which no agent ever specializes ($V_{30}^{low} - V_{20}^{low} < C_3$). A pessimistic society may then be stuck on a Pareto-inferior equilibrium path if the training costs are sufficiently high.

This source of equilibria multiplicity is isomorphic to that identified in [Acemoglu, 1997]. In his model, Acemoglu shows how complementary investments of workers and firms may not be undertaken due to self-fulfilling pessimism; moreover, he demonstrates that the possibility of delay only exacerbates the multiplicity problem. Similarly to his results, my model can also exhibit delayed specialization as an inferior equilibrium path. A policy to overcome this coordination failure, e.g. a short-lived subsidy for training, will be beneficial.

4.7.2 Other wage bargaining specifications

Instead of the equal split of the output each period, wage-setting behavior could be modeled in a more complex way. While I do not endeavour this extension here explicitly, I can still envision some of additional effects introduced. What many of the widely-used bargaining solutions imply is that an increase in the reservation value (impasse point) of a party leads to an increase in her/his share of the surplus. Without formalizing this, I study what happens if the threat point of type 1 workers increases.

Proposition 8 *If an increase in the threat point of type 1 leads to improvement in bargaining position in a match with type 2 and reduction in* w_{21}*, this causes earlier entry of type 3 specialists and earlier decision of type 2 agents not to match with type 1. Thus, an improvement in the bargaining position of type 1 leads to an earlier specialization takeoff, earlier isolation of type 1 workers, and their rising unemployment.*

Proof. Apply Proposition 2.□

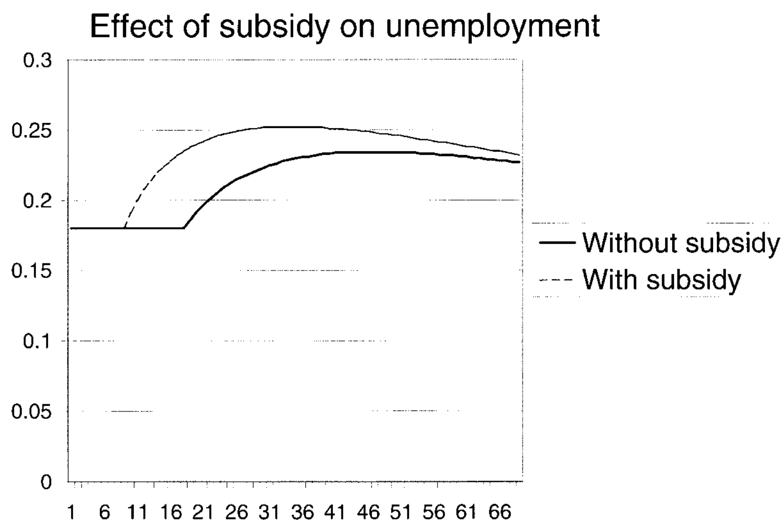

Figure 9:

Consider what happens when government subsidizes obsolete matches $(1,1)$. This increases wage w_{11} and strengthens bargaining position of type 1 workers in other matches. Consequently, workers of type 2 receive lower remuneration in $(1,2)$ matches, while type 3s are unaffected. Hence, specialization becomes relatively more attractive.

Subsidization causes output per capita gains due to earlier specialization but higher transitory unemployment (See Fig. 9). Section 4.8 further discusses the policy potential of this paradoxical implication of subsidization.

4.7.3 Other matching specifications

The random matching assumption is clearly extreme. I can weaken it by allowing for semi-directed search [Acemoglu, 1996b]. As long as there is a positive degree of randomness in the matching, there is a potential for aggregate pecuniary externality to overturn productive efficiency of type 3s.

I could also get the same results under an alternative matching technology. Search is completely directed (workers can form, say, 2 or 3 distinct markets), but the probability of finding a match depends on the *number* of other workers searching in the same pool (IRS matching technology). While this last assumption is usually overturned empirically in developed country contexts, there is evidence that increasing returns play considerable role in transition economies [Münich, Svejnar and Terrell, 1998]. It could be that labor market infrastructure grows locally with the market size (and markets in developed countries represent steady states of this process). Under this alternative specification, the qualitative results are unchanged.

4.8 Some policy implications

As implied by section 4.6, the best case for policy in my model is to try to remedy for the root causes of coordination failure and subsidize specialization.

A policy that makes vocational education relatively cheaper is desirable (cf. Proposition 7 regarding a reduction in C_3), as well as the investment in labor market infrastructure (cf. Proposition 6 regarding changes in m).

A paradoxical implication of proposition 8 is that government sometimes can speed up specialization, and therefore, growth, by subsidizing obsolete matches $(1,1)$. According to proposition 8, this increases wage demands of type 1 workers, pushing down wages of type 2 workers, and making specialization more attractive. One needs to be cautious, however, because this effect is only valid for subsidies that are not large enough. For sufficiently high wage w_{11} neither type 2 nor type 3 training are optimal. Too high a subsidy precludes the adjustment at all. In addition, the subsidy speeds up the isolation of the obsolete specialists, which is undesirable on equity grounds.

Under the "infinite life" interpretation that focuses on mid-life transitions, model has a somewhat unusual implication that better education facilities might under some circumstances slow down growth. This would happen if the incentive effect of the change in education rates dominates the speed effect. As opportunity costs of not retraining decrease, individual workers become more conservative and postpone specialization. However, this does not imply that governments should slash plans to build vocational schools. The efficiency theorem is valid under constant ρ. No efficiency result is available for varying ρ.

5 Facts about labor markets in transition

5.1 Enrollments

It is difficult to provide direct empirical evidence on changes in the aggregate distribution of skills, because the degree of specialization is unobservable or at least very roughly proxied by the data from standard labor force surveys. Further, the somewhat vague distinction in reality between the amount of human capital, and the degree of specialization compounds the interpretation of the facts.

The indicator that perhaps best gauges the special/general training is the ratio of enrollments in vocational and general secondary training, because both might reasonably be described as providing the same *amount* of human capital. Boeri [1999] shows that after reforms started, the composition of these enrollments has changed dramatically across CEE countries (see Fig. 10). Enrollments were higher in general secondary schools and lower in vocational schools in 1996 compared with 1989. These changes in enrollments were demand driven, as evidenced by the figures on the supply of education. In 1994, vocational and technical schools still accounted for 84% of secondary education in the Czech Republic, and 73% in Hungary, compared to 53% OECD average [Boeri et al., 1998].

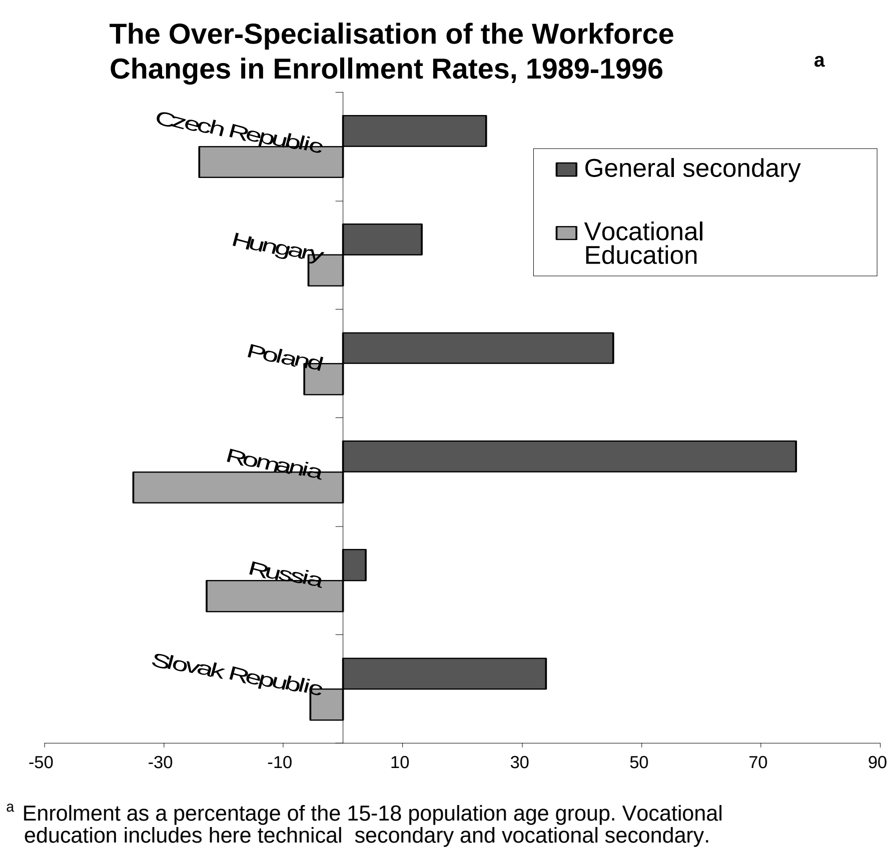

Figure 10: Source: [Boeri, 1999].

In addition to enrollment rates, there is limited evidence on retraining programs for the unemployed. In Czech Republic, starting in 1992, general rather than specific training made up the bulk of the courses organized by labor offices [Ham, Svejnar and Terrell, 1995].

5.2 Education and experience profiles in earnings

Rutkowski [1998a] documents a substantial increase in education premia across most Central European countries. However, a crucial fact is that vocational education ceased to have any effect on earnings [Flanagan, 1998; Rutkowski, 1998a]. In the table below, selected estimates from the log-earnings regressions are presented for Poland, Czech Republic, and Estonia. All of them document flattening of the experience profiles, as skills accumulated under the old system are of little market value.

Experience profiles in log-earnings, selected regressors, CEE countries

	1987	1988	1989	1992	1993	1994	1995
POLAND							
Exp	0.031			0.023	0.021		
Exp^2	−0.056			−0.035	−0.03		
Private exp				0.006	0.005		
Private exp^2				−0.017	−0.039		
CZECH REPUBLIC							
Men							
Years of schooling		0.04		0.053			
Exp		0.046		0.032			
Exp^2		−0.081		−0.066			
Women							
Years of schooling		0.057		0.067			
Exp		0.031		0.016			
Exp^2		−0.051		−0.025			
ESTONIA							
Secondary education			0.081	0.217	0.184	0.234	0.250
Exp			0.139	0.034	0.042	0.055	0.052
Exp^2			−0.027	−0.008	−0.015	−0.02	−0.02

Source: [Rutkowski, 1996], [Večerník, 1995], [Noorkoiv, Orazem, Puur and Vodopivec, 1998].

5.3 Unemployment proportions and incidence

Consistently with predictions of my model, Coricelli, Hegeemejer and Rybinski [1995] find that the largest group in unemployment pool are workers with basic and secondary vocational education. In the third quarter of 1992, for example, the former accounted for 38.2% and 21.8% of unemployment, respectively, while people with secondary general education only for 7.4%. Coricelli et al. [1995] attribute this to shortcomings of Polish educational system, in which vocational schools were closely tied with specific enterprises or branches.

Orazem and Vodopivec [1997] show that in Slovenia workers with vocational skills were significantly more likely to be unemployed. Exit rates from unemployment also fell with age. Their findings are summarized in the table below:

Difference from the baseline probability of finding a job if unemployed. Slovenia.

Experience	Pre-transition	Transition
Less than 3 years	−2.9	2.2
3-5 years	baseline	baseline
5-10 years	4.8	0.7
10-15 years	5.1	0.2
15-20 years	8.0	-0.3
More than 20 years	6.4	5.6

Source: [Orazem and Vodopivec, 1997].

Rutkowski [1998b] documents that narrow vocational skills increase the risk of joblessness in Poland. Cazes and Scarpetta [1998] estimate hazard rates of exit to jobs among registered unemployed in two regions (voivodships) of Poland, Warsaw and Ciechanov. In both single and multiple destination partial hazards econometric models they find that general secondary education has positive and significant effect on exit rates, while vocational training has either insignificant or much lower effect:

Hazard rates of exit to job from unemployment

	Warsaw		Ciechanov	
	1990-1991	1991-1992	1990-1991	1991-1992
Secondary education	−0.007	0.2813*	0.275	0.7141*
Vocational education	0.3255*	0.0806	0.521	0.0636
Multiple destination model				
Secondary education	0.0678	0.8902*		1.1255*
Vocational education	0.2665	0.4377*		0.1074

Source: [Cazes and Scarpetta, 1998]. Asterisks denote statistically significant estimates.

Boeri et al. [1998] show that vocational school graduates in Hungary in 1994 fared worse in terms of the probability of finding a job than students with an equivalent number of years in general secondary schools:

Probability of leaving insured unemployment for a job.
Logit estimates for workers aged 22-26. April 1994, Hungary.

	Odds ratio
Vocational secondary	1.3923
General secondary	1.7476

Source: [Boeri et al., 1998].

Taken together, evidence is broadly consistent with the prediction of my theory that investment in general human capital should be attractive in the beginning phase of transition. I now discuss East German transition experience that contrasts with that in the rest of Central and Eastern Europe.

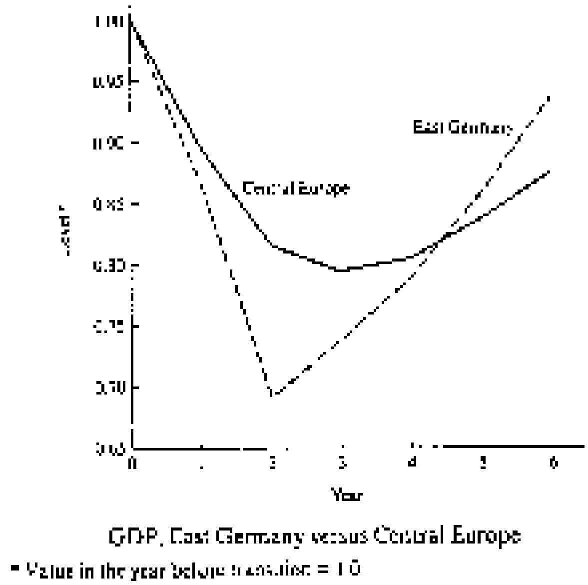

Figure 11: Source: [Blanchard, 1997]

5.4 Unemployment experience of East Germany

There is more to transition than just the adjustment of the human capital stock. Disorganization [Blanchard and Kremer, 1997] definitely explains a significant part of the U-shaped output path. However, there are differences across transition economies that cannot be readily explained in the disorganization framework. For instance, one would expect East Germany to have fewer problems with enforcing contracts than its CEE neighbors, given the powerful influence of the West German body of law, size of direct investment, and the windfall gain in infrastructure capital. But East Germany has experienced the worst unemployment (employment fell by 35%) among the Central and Eastern European countries, and the sharpest output fall, as well as the sharpest recovery (see Fig. 11). Blanchard [1997] suggests that union pressures on wages account for this phenomenon. However, Hunt [1999] finds little support for this hypothesis, as wage growth appears to be similar at growing, shrinking, and stable firms. Moreover, as I document below, wages continued to grow at high rates for the rest of the 90s with no further increase in unemployment.

My model offers an alternative explanation for the size of East German employment fall. The unification enlarged the effective labor market for East Germans. Since this new market was dominated by specialists with modern skills, specialization became optimal at the very beginning. As implied by the model, though, early specialization goes hand in hand with high unemployment.

If this hypothesis is correct, then one would observe a U-shaped pattern in relative returns to vocational vs. general education. I study this implication in the next section.

5.5 Returns to vocational training in East German territories after the unification

Krueger and Pischke [1995] summarize evidence on the earnings distributions in East and West Germany before and immediately after transition. They show that experience profiles in East Germany are very flat like in the rest of CEE. However, East Germany is the only formerly planned economy, where returns to education have dropped in the aftermath of transition. Krueger and Pischke also discuss evidence on job commuters, those East Germans who found jobs in the western lands. The evidence shows that their monthly wages are 83% higher than wages of the workers residing in the East, after controlling for education, experience, and other labor market characteristics.

Hunt [1999] discusses the determinants of wage growth for East German workers. Using median regressions, she finds positive effect on wage growth of secondary education and negative effect of vocational training in 1990-1991. Her evidence for the subsequent (1991-1996) wage growth, is inconclusive.

Determinants of wage growth in German territories, relative to basic apprenticeship

	East, 1990-91	East, 1991-96	West, 1990-96
General schooling	0.113	0.024	-0.003
	(0.062)	(0.021)	(0.003)
Vocational training	-0.032	0.000	-0.003
	(0.014)	(0.006)	(0.002)

Source: [Hunt, 1999]. Standard error in parentheses.

Hunt's [1999] estimates for the subsequent (1991-1996) wage growth, on balance, are indeterminate. Coefficients on either types of schooling alternate between positive and negative across various specifications, and for the most part are insignificant. Hunt's [1999] estimation procedure lumps together old and young workers whose skills are presumably very different. It also does not correct for potential correlation of individual fixed effects with observable characteristics.

Accounting for these shortcomings, I present evidence below that the premium to vocational training relative to general training has been positive and growing for young workers since 1994, in contrast to what has happened in other CEE countries.

For estimation, I use data from the German Socio-Economic Panel. This longitudinal household survey has been conducted in West Germany since 1984; the sample was extended to include East German respondents in 1990. I estimate standard log-earnings regressions for both East and West subsamples for years 1992-1997.

I focus only on respondents with secondary-level general and vocational training[8]. In addition I restrict my sample to young workers born on or after 1967, to concentrate on returns to relatively new vintages of human capital.

To remain in the sample, East German workers must work in the East (approximately 7% of young respondents residing in GDR in 1990 have moved or are commuting to the

[8] They might reasonably be described as having similar investment in human capital in terms of time.

West), and similarly for West German workers. The definitions of the variables are collected in Table 1 below.

Variable definitions

Variable	Description
$\log W$	log gross earnings in the month before the interview
$FEMALE$	female dummy
VOC	dummy = 1 if respondent has vocational secondary education; = 0 if general secondary education
EXP	experience
δ_{st}	year dummy = 1 if current year $t = s$, = 0 otherwise

Table 1

10.2% of the Eastern sample members and 15.8% of Western sample have vocational training over the length of the panel. The specification to be estimated is

$$\begin{aligned} \ln W_{it} &= c + \beta_0 FEMALE_i + \beta_1 VOC_i + \beta_2 EXP_{it} + \beta_3 (VOC_i * EXP_{it}) \\ &\quad + \sum_s \alpha_s \delta_{st} + \sum_s \gamma_s (VOC_i * \delta_{st}) \end{aligned} \tag{11}$$

I estimate equation (11) by fixed effects regression. The results are presented in Table 2 below. Interaction terms of VOC dummy with year dummies are denoted by $VOC9X$. These terms measure the premium to having vocational rather than general schooling in a given year. The omitted category is secondary education; the omitted year is 1992.

Log-earnings regression for East and West German lands

	East	West
VOC	1.708* (.564)	−.085 (.321)
EXP	.100* (.010)	.045* (.007)
$VOC * EXP$	−.224* (.078)	.038 (.051)
$VOC93$	.235 (.170)	−.032 (.140)
$VOC94$	.535* (.2181533)	−.113 (.181)
$VOC95$	.646* (.271)	−.144 (.226)
$VOC96$	.952* (.320)	−.114 (.276)
$VOC97$	1.197* (.378)	−.165 (.321)

Fixed effects estimation with year dummies and a female dummy, base year is 1992; standard errors in parentheses; asterisks indicate coefficients significant at 2% confidence level

Table 2

The results demonstrate that relative returns to vocational schooling have *grown* in the East during years 1992-1997. Wald equality tests further confirm significance of this growth across years. This trend cannot be attributed to the education-neutral time trend, or a trend common to both Eastern and Western German territories. In fact, in the West, there is no significant premium (positive or negative) associated with vocational training.

The returns to experience are positive for workers trained in general schools, and negative for workers with vocational skills. This correlation is clearly driven by the negative effect of vocational schooling acquired under the old system.

Taken together, the evidence on East Germany is broadly supportive of the pattern of adjustment predicted by my theory. To make this statement stronger requires a more detailed empirical study, which I do not endeavor here.

6 Conclusion

The paper outlined a model of human capital adjustment characterized by inertia due to search frictions. The model has important implications for the debate about education policy. [Fan, Overland and Spagat, 1997] worry about the consequences of delayed restructuring for intergenerational human capital transmission. They interpret recent Russian evidence as suggesting that new labor force entrants are not as skilled as before. That leads them to conclude that the quality of labor force is deteriorating and the resumption of growth is seriously threatened. According to my model, this need not be the case. People investing in human capital weigh advantages and drawbacks of narrow specialization and may decide instead to acquire broad specialization. Not only it is optimal for them to do so, it also benefits the society, because the faster the critical mass of Jacks of all trades is accumulated, the faster the entrance of new dab hands will begin. It is therefore counterproductive to support those fields of education that are under-represented since they are likely to increase mismatch and delay the process of adjustment.

The paper further presented cross-coutry evidence on East Germany and other CEE countries, that is indicative of the effects implied by the theory. East Germany can be regarded as an outlier with the speediest adjustment but highest transitory unemployment; this fact can be explained by my model.

At the same time the Czech Republic, sometimes cited among the most successful transition economies in terms of unemployment, may serve as an example of a country whose growth is still held back by inertia. Boeri [1999] presents a theory of hysteresis that explains the lowest unemployment rate among all CEE countries by lower unemployment benefits at the start of transition. But Czech output also grew at a slower rate than the neighbors'. My model can account for these discrepancies. If unemployment cushion is absent, transition of the economy is delayed. Indeed, Münich et al. [1998] suggest that delayed restructuring may be responsible for lower unemployement in the Czech Republic.

None of the existing empirical studies explore systematically the differences and changes in labor market prospects of workers with general and special skills. My model provides a framework to analyze these questions in a structured way.

Lastly, predictions of the model may also be relevant in more general contexts, in any country experiencing an adjustment of its labor force to a large technology, organization, or demand shock. In the context of persistent TFP differentials across countries, some explanations [Acemoglu and Zilibotti, 1998] have emphasized very dissimilar endowments of skills across countries. My model can help explain why these distributions exhibit lack of convergence. It can also account for the U-shaped Kuznets curve [Kuznets, 1955] both across countries and over time.

A Analytical framework and proofs of propositions

Appendices contain the formal part of the analysis. Appendix A builds the stochastic framework for the analysis of labor market transitions and proves propositions of section 4.5. Appendix B analyzes efficiency of adjustment, focusing on the optimality of specialization threshold. Appendix C contains proofs of auxiliary results.

A.1 Transitions across the employment states, the generator matrix, and the value function

There are a total of 12 possible Markov states in which an agent can dwell in any moment of time: three possible own types times four possible employment states:

type 1	×	unemployed
type 2		working in pair with type 1
type 3		working in pair with type 2
		working in pair with type 3

Deferring for a moment the possibility of changing own type, let us describe the transition matrix for a worker of a *given type* (thus focussing on transitions between just 4 states). The dynamic structure is a continuous-time Markov process. One may note that this is a generalized Poisson process with time-varying parameters (because q_j change over time). The previous job history does not matter for the prospects of finding or maintaining a working match, only the current state matters. We will index the state by (i, j), where $i = 1, 2, 3$ and $j = 0, 1, 2, 3$, with i standing for own type, and j for employment state. For example $(2, 0)$ denotes an unemployed worker of type 2.

Markov process is specified by its transition function $P(s, (i, j), t, (i', j'))$. Since we have a finite number of states, this is a matrix 4×4. To compute it, we first write down the generator matrix[9] $G_{ii}(t)$, specifying transition rates between the states:

$$G_{ii}(t) = \begin{pmatrix} -m(q_1 + q_2 + q_3) & mq_1 & mq_2 & mq_3 \\ \delta & -\delta & 0 & 0 \\ \delta & 0 & -\delta & 0 \\ \delta & 0 & 0 & -\delta \end{pmatrix}. \tag{12}$$

We may simplify this by noting that $q_1 + q_2 + q_3 \equiv 1$. $G(t)$ varies over time because q_j do.

In the generator matrix, arrival of a training opportunity is accounted for by a transition rate ρ. *Where* does the transition occur is the decision variable of the agent, thus some elements of G are the policy function. The general form of the matrix G is

$$G = \begin{pmatrix} G_{11} & G_{12} & G_{13} \\ G_{21} & G_{22} & G_{23} \\ G_{31} & G_{32} & G_{33} \end{pmatrix}$$

This partitioned form gives meaning to expression (12). The structure of matrix G is best

[9] The meaning of the subscript ii will become clear shortly. This indicates that G_{ii} is a submatrix of a larger matrix 12×12 describing transitions between all 12 states.

illustrated by an example:

$$\left(\begin{array}{cccc|cccc|cccc}
-(m+\rho) & mq_1 & mq_2 & mq_3 & \rho\lambda & 0 & 0 & 0 & \rho(1-\lambda) & 0 & 0 & 0 \\
\delta & -\delta & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\delta & 0 & -\delta & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\delta & 0 & 0 & -\delta & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
 & & & & -(m+\rho) & mq_1 & mq_2 & mq_3 & \rho & 0 & 0 & 0 \\
 & \multicolumn{2}{c}{0} & & \delta & -\delta & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & \delta & 0 & -\delta & 0 & 0 & 0 & 0 & 0 \\
 & & & & \delta & 0 & 0 & -\delta & 0 & 0 & 0 & 0 \\
\hline
 & & & & & & & & -m & mq_1 & mq_2 & mq_3 \\
 & \multicolumn{2}{c}{0} & & & \multicolumn{2}{c}{0} & & \delta & -\delta & 0 & 0 \\
 & & & & & & & & \delta & 0 & -\delta & 0 \\
 & & & & & & & & \delta & 0 & 0 & -\delta
\end{array}\right) \tag{13}$$

Matrices G_{21}, G_{31}, G_{32}, although not restricted to zero in the model, will be so on equilibrium path, because no worker will desire to downgrade one's own type. All zero elements except the top-left one in matrix G_{23} indicate that there are no potential transitions between type 2 and type 3 other than through unemployment pools. $\rho\lambda$ and $\rho(1-\lambda)$ in top left elements of G_{12}, G_{13} indicate that in state $(1,0)$, when a training opportunity arrives, worker randomizes between choosing type 2 and type 3 with probabilities λ, $1-\lambda$. Again, G in (13) is just a particular example of how this matrix may look, given agents' choices. This sample matrix also says there is no discrimination in matching decisions. Everyone matches with either type as implied by non-zero transition rates into all states (i,j).

The transition matrix is then given by

$$P(s,t) = \exp\left(\int_s^t G(\tau)\, d\tau\right).$$

If G did not vary, this would simplify to $P(s,t) = e^{(t-s)G}$, where matrix exponential is defined as converging series $e^A = I + A + \frac{A^2}{2} + \frac{A^3}{3!} + \ldots$

Let $\mathbf{p}(s)$ be a row vector giving the probability distribution over employment states at time s. We start the process at time 0 from the state $(1,1)$ (all workers are of type 1 and there is full employment). By Markov property, $\mathbf{p}(s) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} P(0,s)$. We denote its elements by $p_{(i,j)}(s)$ which correspond to the probability of state (i,j).

As implied by the form of the generator matrix, I will use the following convention when referencing the elements of the derived transition matrix: $P_{(i,j)\to(k,n)}(s,t)$ specifies the element with row index $(i-1)*3+j$ and column index $(k-1)*3+n$ (probability of transition from state (i,j) in period s to state (k,n) in period t).

Finally, I endow workers with an additional choice variable: ϕ_{ij}. If $\phi_{ij}=1$, worker of type i agrees to stay in productive partnership with type j worker. If $\phi_{ij}=0$, we say that type i decides not to match with j (decreasing, obviously, one's probability of finding a match). Obviously, what is important from the viewpoint of an agent, is not who decides to break a match, oneself or the other party in the partnership. What matters is whether someone breaks the match at all. Hence it is the product $\phi_{ij}\phi_{ji} \equiv \psi_{ij}$ that is relevant, and the general form of matrix G_{ii} is

$$G_{ii} = \begin{pmatrix} -m\sum_{j=1}^{3} q_j\psi_{ij} & mq_1\psi_{i1} & mq_2\psi_{i2} & mq_3\psi_{i3} \\ \delta\psi_{i1} & -\delta\psi_{i1} & 0 & 0 \\ \delta\psi_{i2} & 0 & -\delta\psi_{i2} & 0 \\ \delta\psi_{i3} & 0 & 0 & -\delta\psi_{i3} \end{pmatrix}$$

We assume that when a type i worker sets one's $\phi_{ij} = 0$ while still employed with type j, she/he transits into the unemployment state instantly at will.

For example a type 2 worker who avoids matches with type 1 worker and is at the same time shunned by type 3 workers has

$$G_{ii} = \begin{pmatrix} -mq_2 & 0 & mq_2 & 0 \\ 0 & 0 & 0 & 0 \\ \delta & 0 & -\delta & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

It is now time to specify payoffs received by agents along their working history. I denote by w_{ij} the wage paid to i in a match (i, j). In what follows I will assume for simplicity that a pair of workers simply split their output evenly:

$$w_{ij} = w_{ji} = y_{ij}/2.$$

z denotes the flow return to unemployed agents. For uniformity I will sometimes use alternative notation $w_{i0} = z$. This allows to define the wage vector:

$$\mathbf{w} = \begin{pmatrix} w_{10} & w_{11} & w_{12} & w_{13}| & w_{20} & w_{21} & w_{22} & w_{23}| & w_{30} & w_{31} & w_{32} & w_{33} \end{pmatrix}'.$$

Concluding this section I write down the expression for the discounted expected utility (value function) of an agent of type i working with type j (or unemployed if $j = 0$) at time s, $V_{ij}(s)$. Obviously, it is a function of the state (i, j). The expected instantaneous utility at moment t given the initial conditions is given by $\sum_{(k,n)} P(s, (i, j), t, (k, n)) w_{kn}$. Integrating this over time gives

$$V_{ij}(s) = \int_s^{+\infty} e^{-r(t-s)} \sum_{(k,n)} P(s, (i, j), t, (k, n)) w_{kn} dt. \quad (14)$$

On some occasions in the paper I will also use notation $\mathbf{V}(t)$ for the value *vector* (of dimension 12):

$$\mathbf{V}(s) = \int_s^{+\infty} e^{-r(t-s)} P(s, t) \mathbf{w} dt.$$

A.2 Optimal decision thresholds

The preceding section has specified that agents only have freedom in choosing their type when a training opportunity arrives and in accepting or passing other types of agents when matched together. Let us rewrite (14) in vector form:

$$\begin{aligned} V_{ij}(s) &= \int_s^{+\infty} e^{-r(t-s)} \sum_{(k,n)} \left[\exp\left(\int_s^t G(\tau)\, d\tau \right) \right]_{(k,n)} w_{kn} dt \\ &= \int_s^{+\infty} e^{-r(t-s)} \mathbf{d}_{ij} \exp\left(\int_s^t G(\tau)\, d\tau \right) \mathbf{w} dt, \end{aligned} \tag{15}$$

where $\mathbf{d}_{ij}$ is a unit vector that has all zero coordinates except coordinate $(i-1)*3+j$. Clearly, the optimal decision of an individual in the model depends on the actions of the other agents (reflected in the generator matrix G); however, the individual should only be concerned with the aggregates. First, the proportions of types in unemployment q_1, q_2, q_3 depend on the *past* actions of the other agents. Second, ψ_{ij} depend on the others' *current* matching choice. For all agents of type i, optimal choice ϕ_{ij} is the same and unique. Thus, q_1, q_2, q_3 and ϕ_{ij} are the aggregates that influence the utility of the agent.

A.3 Equilibrium thresholds

The stochastic process that governs transitions between Markov states for each agent is generalized Poisson process with time-varying parameter. The agent has control over some of the parameters of the transition matrix (for example, by deciding to avoid working with workers of type 1, type 2 worker changes the flow rates into the state "working with agent of type 1"). This control is essentially binary – at each point in time, if an opportunity arrives, an agent is free to choose whether to take or not some action.

The most interesting threshold is the specialization threshold, $V_{20}(s) = V_{30}(s)$. But what are the policies (generator matrices) associated with these utilities? For this to be a threshold, the generator matrix should in both cases prescribe that upon arrival of a training opportunity, agent chooses type 3. Thus,

$$G(s) = \left[\begin{array}{c|cccccccc} G_{11} & & G_{12} & & & & G_{13} & & \\ \hline & -m-\rho & mq_1 & mq_2 & mq_3 & \rho & & & \\ & \lambda & -\lambda & 0 & 0 & & 0 & & \\ & \lambda & 0 & -\lambda & 0 & & & 0 & \\ 0 & \lambda & 0 & 0 & -\lambda & & & & 0 \\ & & & & & -m(q_2+q_3) & 0 & mq_2 & mq_3 \\ & & & & & \lambda & -\lambda & 0 & 0 \\ & & & & & \lambda & 0 & -\lambda & 0 \\ & & & & & \lambda & 0 & 0 & -\lambda \end{array} \right]. \tag{16}$$

In the remainder of this section we will ignore the irrelevant blocks to the north and to the west, and concentrate on the 8×8 south-eastern block of this matrix.

Expressing $V_{20}(s)$ and $V_{30}(s)$ as in (15) and using matrix $G(s)$ defined above we obtain

$$V_{20}(s) = V_{30}(s) \Leftrightarrow \int_s^{+\infty} e^{-r(t-s)} \left(\mathbf{d}_{2,0} - \mathbf{d}_{3,0}\right) \exp\left(\int_s^t G(\tau)\, d\tau\right) \mathbf{w} dt. \tag{17}$$

Here $\mathbf{d}_{2,0} - \mathbf{d}_{3,0} = \begin{pmatrix} 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \end{pmatrix}$.

In equilibrium all agents' thresholds coincide because they are symmetric, and in addition $q_3 = 0$, $q_1 + q_2 = 1$ (by definition of the threshold). Thus

$$G(s) = \begin{bmatrix} -m-\rho & mq_1 & mq_2 & 0 & \rho & & & \\ \lambda & -\lambda & 0 & 0 & & 0 & & \\ \lambda & 0 & -\lambda & 0 & & & 0 & \\ \lambda & 0 & 0 & -\lambda & & & & 0 \\ & & & & -mq_2 & 0 & mq_2 & 0 \\ & & & & \lambda & -\lambda & 0 & 0 \\ & & & & \lambda & 0 & -\lambda & 0 \\ & & & & \lambda & 0 & 0 & -\lambda \end{bmatrix} \tag{18}$$

at equilibrium threshold. I will denote the solution to (17), (18) by s_0:

$$\int_{s_0}^{+\infty} e^{-r(t-s)} \left(\mathbf{d}_{2,0} - \mathbf{d}_{3,0}\right) P(s_0, t)\, \mathbf{w} dt = 0. \tag{19}$$

I study the following other thresholds: $s_{obs} = \inf\{t \mid \phi_{21}(t) = 0\}$ (the *obsolescence* threshold), $s_{negobs} = \sup\{t \mid \phi_{21}(t) = 0\}$, $s_{seg} = \inf\{t \mid \phi_{32}(t) = 0\}$ (the *segregation* threshold).

A.3.1 Obsolescence thresholds

Workers of type 2 will prefer to forego a match with type 1 if being unemployed and searching gives them higher expected utility: $V_{21} < V_{20}$. Thus, the moment s_{obs} satisfies

$$\int_{s_{obs}}^{+\infty} e^{-r(t-s)} \left(\mathbf{d}_{2,1} - \mathbf{d}_{2,0}\right) P(s_0, t)\, \mathbf{w} dt = 0. \tag{20}$$

Workers of type 2 may later revert to matching with type 1, and the moment they do so, s_{negobs}, satisfies the same condition (7).

A.3.2 Segregation threshold

Moment s_{seg} is when the new specialists (type 3) start avoiding matches with type 2. It is given by the equation $V_{32} = V_{30}$, or

$$\int_{s_{seg}}^{+\infty} e^{-r(t-s)} \left(\mathbf{d}_{3,2} - \mathbf{d}_{3,0}\right) P(s_{obs}, t)\, \mathbf{w} dt = 0. \tag{21}$$

A.4 Comparative dynamics in detail

Propositions below are proved in the partial equilibrium setting, i.e. under the assumption that other thresholds do not change. Theorem 1 then proves that the thresholds are supermodular [Topkis, 1998]. Supermodularity ensures that signs of all partial equilibrium effects are preserved in general equilibrium. Therefore, all propositions below hold unconditionally.

Proposition 2. *A decrease in w_{21} leads to earlier specialization takeoff, earlier and longer isolation of type 1 workers. Although there is no* direct *effect of w_{21} on the segregation threshold, the general equilibrium effect is negative.*

Proof. Differentiate (19) w.r.t. w_{21}:

$$\begin{aligned}\frac{\partial}{\partial w_{21}}\int_{s_0}^{+\infty} e^{-r(t-s)}\left(\mathbf{d}_{2,0}-\mathbf{d}_{3,0}\right)P\left(s_0,t\right)\mathbf{w}dt &= \int_{s_0}^{+\infty} e^{-r(t-s)}\left(\mathbf{d}_{2,0}-\mathbf{d}_{3,0}\right)P\left(s_0,t\right)\frac{\partial}{\partial w_{21}}\mathbf{w}dt \\ &= \int_{s_0}^{+\infty} e^{-r(t-s)}\left(P_{(2,0)\to(2,1)}-P_{(3,0)\to(2,1)}\right)dt \\ &= \int_{s_0}^{+\infty} e^{-r(t-s)}P_{(2,0)\to(2,1)}\left(s_0,t\right)dt>0,\end{aligned}$$

since $P_{(2,0)\to(2,1)}\geq 0$, $P_{(3,0)\to(2,1)}\equiv 0$[10]. Therefore, an increase in w_{21} will lead to a shift in the relative attractiveness schedule $(V_{20}-V_{30})$ upward, at any given q_2 decreasing the incentive to specialize. On the other hand, a *reduction* in w_{21} has an opposite effect, so specialization takes off earlier.

Now consider the effects on the obsolescence thresholds. Differentiate (7) w.r.t. w_{21}:

$$\begin{aligned}&\frac{\partial}{\partial w_{21}}\int_{s_{obs}}^{+\infty} e^{-r(t-s)}\left(\mathbf{d}_{2,1}-\mathbf{d}_{2,0}\right)P\left(s_{obs},t\right)\mathbf{w}dt \\ =&\int_{s_{obs}}^{+\infty} e^{-r(t-s)}\left(P_{(2,1)\to(2,1)}\left(s_{obs},t\right)-P_{(2,0)\to(2,1)}\left(s_{obs},t\right)\right)dt\end{aligned}$$

By the general property of Markov chains with a stationary distribution, $P_{(2,1)\to(2,1)}\left(\cdot,t\right)$ is decreasing in t from 1, while $P_{(2,0)\to(2,1)}\left(\cdot,t\right)$ is increasing from 0, so the derivative is positive.

There is no partial equilibrium effect on the segregation threshold s_{seg}, as the derivative of condition (21) w.r.t. w_{21} is zero.

□

Proposition 3. *An increase in unemployment benefit z leads to earlier specialization, earlier isolation of type 1, and earlier segregation of type 3 workers.*

Proof. Differentiate (19) w.r.t. z:

$$\begin{aligned}&\frac{\partial}{\partial z}\int_{s_0}^{+\infty} e^{-r(t-s)}\left(\mathbf{d}_{2,0}-\mathbf{d}_{3,0}\right)P\left(s_0,t\right)\mathbf{w}dt = \int_{s_0}^{+\infty} e^{-r(t-s)}\left(\mathbf{d}_{2,0}-\mathbf{d}_{3,0}\right)P\left(s_0,t\right)\frac{\partial}{\partial z}\mathbf{w}dt \\ =&\int_{s_0}^{+\infty} e^{-r(t-s)}\left(P_{(2,0)\to(2,0)}+P_{(2,0)\to(3,0)}-P_{(3,0)\to(3,0)}\right)dt<0\end{aligned}$$

[10] Since probability of downgrading from type 3 to type 2 is zero.

In order to see why the last inequality holds, notice that $\left(P_{(2,0)\to(2,0)} + P_{(2,0)\to(3,0)}\right)$ is the probability that an unemployed worker of type 2 at time s_0 will remain unemployed (of any type). These odds are lower than for type 3 unemployed worker, because the flow rate to jobs is higher for type 2 as long as $q_1 > 0$. More formally, recall that the transition probabilities $\left(P_{(2,0)\to(2,0)} + P_{(2,0)\to(3,0)}\right)$ and $P_{(3,0)\to(3,0)}$ are driven by the continuous Markov process. In the generator matrix state $(2,0)$ has higher absolute outflow rate from unemployment m compared to $m(q_2 + q_3)$ in state $(3,0)$ (cf. matrix in (16)). Matrix P is obtained as $\exp\left(\int_s^t G d\tau\right)$. By the monotonicity of the matrix exponential w.r.t. elements, $P_{(2,0)\to(2,0)} + P_{(2,0)\to(3,0)} < P_{(3,0)\to(3,0)}$.

To determine the effect on the obsolescence thresholds, differentiate (7) w.r.t. z:

$$\begin{aligned}
&\frac{\partial}{\partial z}\int_{s_{obs}}^{+\infty} e^{-r(t-s)}\left(\mathbf{d}_{2,1} - \mathbf{d}_{2,0}\right) P\left(s_{obs}, t\right)\mathbf{w}dt \\
=& \int_{s_{obs}}^{+\infty} e^{-r(t-s)}\left(P_{(2,1)\to(2,0)} - P_{(2,0)\to(2,0)} + P_{(2,1)\to(3,0)} - P_{(2,0)\to(3,0)}\right) dt < 0
\end{aligned}$$

Similarly to the argument above, $P_{(2,1)\to(2,0)} < P_{(2,0)\to(2,0)}$, $P_{(2,1)\to(3,0)} < P_{(2,0)\to(3,0)}$, and so the effect is negative.

To determine the effect on the segregation threshold, differentiate (21) w.r.t. z:

$$\begin{aligned}
&\frac{\partial}{\partial z}\int_{s_{seg}}^{+\infty} e^{-r(t-s)}\left(\mathbf{d}_{3,2} - \mathbf{d}_{3,0}\right) P\left(s_{obs}, t\right)\mathbf{w}dt \\
=& \int_{s_{seg}}^{+\infty} e^{-r(t-s)}\left(P_{(3,2)\to(3,0)} - P_{(3,0)\to(3,0)}\right) dt < 0
\end{aligned}$$

□

Proposition 4. *An increase in training opportunity arrival rate ρ has an ambiguous effect on threshold expressed in terms of q_2. The incentive effect is positive, while the speed effect is negative.*

Proof. For a representative agent, differentiate the threshold condition (19) w.r.t ρ :

$$\frac{\partial}{\partial \rho}\left(V_{20} - V_{30}\right) = \int_{s_0}^{+\infty} e^{-r(t-s)}\left(\mathbf{d}_{2,0} - \mathbf{d}_{3,0}\right)\left[\frac{\partial}{\partial \rho}P\left(s_0, t\right)\right]\mathbf{w}dt \tag{22}$$

$$\begin{aligned}
\frac{\partial}{\partial \rho}P\left(s_0, t\right) &= \frac{\partial}{\partial \rho}\exp\left(\int_{s_0}^{t} G(\tau)\, d\tau\right) = \left(\int_{s_0}^{t}\frac{\partial G(\tau)}{\partial \rho}d\tau\right)\exp\left(\int_{s_0}^{t} G(\tau)\, d\tau\right) \\
&= \left(\int_{s_0}^{t}\frac{\partial G(\tau)}{\partial \rho}d\tau\right) P\left(s_0, t\right)
\end{aligned}$$

$$\frac{\partial G(\tau)}{\partial \rho} = \left(\begin{array}{cccc|cccc}
-1 & m\frac{\partial q_1}{\partial \rho} & m\frac{\partial q_2}{\partial \rho} & m\frac{\partial q_3}{\partial \rho} & & 1 & & \\
& \cdots & & & & & 0.. & \\
& & & 0 & & & 0 & \\
\hline
& & & & m\frac{\partial q_1}{\partial \rho} & 0 & m\frac{\partial q_2}{\partial \rho} & m\frac{\partial q_3}{\partial \rho} \\
& 0 & & & & \cdots & & \\
& & & & & & & 0
\end{array}\right)$$

Premultiplying by $(\mathbf{d}_{2,0}-\mathbf{d}_{3,0})$ gives

$$\begin{aligned}&(\mathbf{d}_{2,0}-\mathbf{d}_{3,0})\frac{\partial G(\tau)}{\partial\rho} \\ = \ & \underbrace{(\mathbf{d}_{3,0}-\mathbf{d}_{2,0})}_{\Gamma}+\underbrace{m\begin{pmatrix} 0 & \frac{\partial q_1}{\partial\rho} & \frac{\partial q_2}{\partial\rho} & \frac{\partial q_3}{\partial\rho} & -\frac{\partial q_1}{\partial\rho} & 0 & -\frac{\partial q_2}{\partial\rho} & -\frac{\partial q_3}{\partial\rho}\end{pmatrix}}_{\Lambda}\end{aligned} \tag{23}$$

Applying lemma 1,

$$\begin{aligned}\frac{\partial}{\partial\rho}(V_{20}-V_{30}) &= \int_{s_0}^{+\infty}e^{-r(t-s_0)}\left(\int_{s_0}^{t}\Gamma d\tau\right)P(s_0,t)\,\mathbf{w}dt+\int_{s_0}^{+\infty}e^{-r(t-s_0)}\left(\int_{s_0}^{t}\Lambda d\tau\right)P(s_0,t)\,\mathbf{w}dt \\ &= \int_{s_0}^{+\infty}e^{-r(t-s_0)}\Gamma\mathbf{V}(t)\,dt+\int_{s_0}^{+\infty}e^{-r(t-s_0)}\Lambda\mathbf{V}(t)\,dt\end{aligned}$$

I distinguish between the incentive effect and the speed effect of a change in ρ. Term Γ comes from varying the individual decision of the agent whether to transit or not into state $(3,0)$ upon arrival of a training opportunity. I call this the incentive effect of a change in ρ. Term Λ appears because the aggregate dynamics of the population depends on ρ. Intuitively, more agents of type 1 become either type 2 before the threshold, or type 3 after the threshold. As a result of higher ρ, transition speeds up. I call this the speed effect of a change in ρ. Strictly speaking, both incentive and speed effects have to do with *incentives* of an agent to specialize, but this distinction allows to decompose the total effect in an intuitive way.

Consider the incentive effect first: $\Gamma=(\mathbf{d}_{3,0}-\mathbf{d}_{2,0})$. Using lemma 1,

$$\begin{aligned}\int_{s_0}^{+\infty}e^{-r(t-s_0)}\Gamma\mathbf{V}(t)\,dt &= \int_{s_0}^{+\infty}e^{-r(t-s_0)}(\mathbf{d}_{3,0}-\mathbf{d}_{2,0})\,\mathbf{V}(t)\,dt \\ &= \int_{s_0}^{+\infty}e^{-r(t-s_0)}(V_{30}-V_{20})\,dt>0\end{aligned}$$

The incentive effect is positive.

Consider now term Λ in (23). Lemma 5 shows that after the threshold $\frac{\partial q_3(t)}{\partial\rho}>0$, $\frac{\partial q_1(t)}{\partial\rho}<0$, $\frac{\partial q_2(t)}{\partial\rho}<0$. By lemma 4, substituting ρ for x,

$$\int_{s_0}^{+\infty}e^{-r(t-s_0)}\Lambda\mathbf{V}(t)\,dt<0$$

The speed effect on the threshold is negative. Notice that the speed effect in itself can be decomposed into two components, effect of the past and of the future. Better *future* prospects due to higher ρ imply lower threshold in terms of q_2, while higher speed of transition *in the past* implies that any fixed threshold will be achieved faster. Using lemma 2, and differentiating (29) w.r.t. ρ gives

$$\left.\frac{\partial q_1}{\partial\rho}\right|_{t<s_0}=-\frac{2q_1\eta}{\xi^2}-\frac{(\xi-\eta)\,e^{\lambda_2 t}-(\eta+\xi)\,e^{\lambda_1 t}}{2\xi\left(e^{\lambda_2 t}-e^{\lambda_1 t}\right)}q_1 t,$$

which is everywhere negative. This is very intuitive, since higher outflow into type 2 reduces the proportion of type 1 in unemployment, q_1, at all times.

We have seen that the incentive and speed effects of higher ρ are of opposite signs. It is impossible to tell a priori which of the effects will dominate. One may still speculate that when the returns to specialization are very high, $(y_{33} - 2y_{23} + y_{22}) \gg 0$, the speed effect is likely to dominate, and the total effect of a change in ρ on (19) is negative, implying lower specialization threshold in terms of q_2. $\square$

Proposition 5. *An increase in job destruction rate δ leads to postponed specialization takeoff: s_0 increases.*

Proof. An increase in δ acts through expectation of longer unemployment spells in the future and slower transition. Similarly to the preceding proof, differentiate the integrand of (19) w.r.t δ and premultiply by $(\mathbf{d}_{2,0} - \mathbf{d}_{3,0})$

$$\begin{aligned} & (\mathbf{d}_{2,0} - \mathbf{d}_{3,0}) \frac{\partial G(\tau)}{\partial \delta} \\ = \; & m \begin{pmatrix} 0 & \frac{\partial q_1}{\partial \delta} & \frac{\partial q_2}{\partial \delta} & \frac{\partial q_3}{\partial \delta} & -\frac{\partial q_1}{\partial \delta} & 0 & -\frac{\partial q_2}{\partial \delta} & -\frac{\partial q_3}{\partial \delta} \end{pmatrix} \end{aligned} \tag{24}$$

Here, the incentive effect is zero. Now by lemma 5, $\frac{\partial q_3(t)}{\partial \delta} < 0$, $\frac{\partial q_1(t)}{\partial \delta} > 0$, $\frac{\partial q_2(t)}{\partial \delta} > 0$ after the threshold. Therefore, by lemma 4, substituting δ for x, the speed effect is positive

$$\int_{s_0}^{+\infty} e^{-r(t-s)} (\mathbf{d}_{2,0} - \mathbf{d}_{3,0}) \frac{\partial G(\tau)}{\partial \delta} \mathbf{V}(t)\, dt > 0$$

An increase in δ shifts the relative attractiveness schedule $(V_{20} - V_{30})$ upward, at any given q_1 decreasing the incentive to specialize. Separating the effect of the past is again straightforward: using lemma 2, and differentiating (29) w.r.t. δ gives

$$\left.\frac{\partial q_1}{\partial \delta}\right|_{t<s_0} = 2q_1 \left(\frac{2(m+\rho)^2 + \delta(m-\rho)}{\delta \xi^2} + \frac{(\xi+\phi)e^{\lambda_1 t} + (\xi-\phi)e^{\lambda_2 t}}{4\xi\left(e^{\lambda_2 t} - e^{\lambda_1 t}\right)} t \right),$$

where $\phi = 2m + \delta - 2\rho$, $\xi > \phi$. A generous sufficient condition for this derivative to be positive is $m > \rho$. Thus the effect of higher δ is always positive.

Proposition 6. *An increase in matching rate m has a negative incentive effect on the specialization threshold to bet on type 3; a sufficient condition for the speed effect to be negative is that half of the unemployed population remains of type 1 at the threshold.*

Proof. Similarly to the preceding proof, differentiate (19) w.r.t m and invoke lemma 1:

$$\int_{s_0}^{+\infty} e^{-r(t-s_0)} (\mathbf{d}_{2,0} - \mathbf{d}_{3,0}) \left(\int_{s_0}^{t} \frac{\partial G(\tau)}{\partial m} d\tau \right) P(s_0, t)\, \mathbf{w} dt = \int_{s_0}^{+\infty} e^{-r(t-s_0)} (\mathbf{d}_{2,0} - \mathbf{d}_{3,0}) \frac{\partial G(t)}{\partial m} \mathbf{V}(t)\, dt$$

$$\begin{aligned} \frac{\partial G(t)}{\partial m} \mathbf{V}(t) &= q_1 V_{21} + q_2 V_{22} + q_3 V_{23} - V_{20} + (q_2 + q_3) V_{30} - q_2 V_{32} - q_3 V_{33} \\ &\quad + m \left(-\frac{\partial q_3}{\partial m} [(V_{33} - V_{30}) - (V_{23} - V_{21})] - \frac{\partial q_2}{\partial m} [(V_{32} - V_{30}) - (V_{22} - V_{21})] \right) \\ &= \Gamma + \Lambda \end{aligned}$$

To show that the incentive effect, gauged by Γ, is positive, recall the Bellman asset value equations for the unemployment states:

$$rV_{20} = z + \sum_j mq_j\psi_{2j}\left(V_{2j} - V_{20}\right) + \dot{V}_{20}, \quad (25)$$

$$rV_{30} = z + \sum_j mq_j\psi_{3j}\left(V_{3j} - V_{30}\right) + \dot{V}_{30}. \quad (26)$$

Hence, $r\left(V_{20} - V_{30}\right) - \left(\dot{V}_{20} - \dot{V}_{30}\right) = \sum_j mq_j\left(\psi_{2j}\left(V_{2j} - V_{20}\right) - \psi_{3j}\left(V_{3j} - V_{30}\right)\right)$. It follows that

$$\begin{aligned} & q_1V_{21} + q_2V_{22} + q_3V_{23} - V_{20} + \left(q_2 + q_3\right)V_{30} - q_2V_{32} - q_3V_{33} \\ = & \frac{r}{m}\left(V_{20} - V_{30}\right) - \frac{1}{m}\left(\dot{V}_{20} - \dot{V}_{30}\right) \end{aligned}$$

The incentive effect is

$$\begin{aligned} \int_{s_0}^{+\infty} e^{-r(t-s_0)}\Gamma dt &= \frac{r}{m}\int_{s_0}^{+\infty} e^{-r(t-s_0)}\left(V_{20} - V_{30}\right)dt - \left[V_{20}\left(s_0\right) - V_{30}\left(s_0\right)\right] \\ &= \frac{r}{m}\int_{s_0}^{+\infty} e^{-r(t-s_0)}\left(V_{20} - V_{30}\right)dt < 0 \end{aligned}$$

The last equality follows because $V_{20}\left(s_0\right) = V_{30}\left(s_0\right)$ by the definition of the threshold s_0. The integral converges because V_{ij} is bounded from above. Summarizing, the incentive effect of a decrease in market frictions leads to earlier specialization.

Let us consider the speed effect. By lemma 5, $\frac{\partial q_3(t)}{\partial m} > 0$, $\frac{\partial q_1(t)}{\partial m} < 0$, $\frac{\partial q_2(t)}{\partial m} < 0$ after the threshold. Therefore, by lemma 4, substituting m for x, the speed effect, Λ, is also negative.

Thus, both the speed and the incentive effects of an increase in m lead to earlier specialization decision. In terms of q_2, the threshold decreases. However, it may take more time to reach this threshold, because, by lemma 2

$$\left.\frac{\partial q_1}{\partial m}\right|_{t<s_0} = q_1\frac{\delta\phi}{\xi^2 m} + \left(2m + \delta\right)t\frac{\left(\xi + \psi\right)e^{\lambda_1 t} + \left(\psi - \xi\right)e^{\lambda_2 t}}{2\xi^2\left(1 - e^{-\frac{1}{2}(2m+\delta)t}\right)} > 0$$

□

Theorem 1 *Thresholds s_0, s_{obs}, s_{seg} are supermodular. A shift in either threshold leads to shifts in other thresholds in the same direction. Therefore, the signs of partial and general equilibrium effects on thresholds s_0, s_{obs}, s_{seg}, coincide.*

I summarize the supermodularity of thresholds by means of table 3. Its elements contain signs of the incentive/speed effects, respectively, of a change in the column threshold on the row threshold.

Effect of:

Effect on:	s_0	s_{obs}	s_{seg}
s_0	*	0/+	+/+
s_{obs}	0/+	*	0/+
s_{seg}	0/+	0/0	*

Table 3

I will demonstrate the sign of the incentive effect of an increase in s_{seg} on s_0. Similarly to appendix B, I perturb the relative attractiveness schedule $(V_{20} - V_{30})$ by increasing s_{seg} by small h, and take the limit of the variation:

$$\begin{aligned}\lim_{h \downarrow 0} \frac{(V_{20} - V_{30}) - \left(\tilde{V}_{20} - \tilde{V}_{30}\right)}{h} &= \int_{s_0}^{+\infty} e^{-r(t-s)} \begin{pmatrix} 0 & 0 & 0 & mq_3 & 0 & 0 & 0 & 0 \end{pmatrix} \vec{\mathbf{V}} dt \\ &= m \int_{s_0}^{+\infty} e^{-r(t-s)} q_3 V_{23} dt > 0\end{aligned}$$

Intuitively, when type 3 workers defer segregation, it makes life of type 2 workers less hazardous, hence specialization is postponed.

I prove formally the signs of the speed effects of a change in s_0 on s_{seg} and s_{obs}. For that I differentiate the threshold conditions w.r.t. s_0. This affects the proportions q_j in the generator matrix:

$$\begin{aligned}&\frac{\partial}{\partial s_0} \int_{s_{seg}}^{+\infty} e^{-r(t-s)} \left(\mathbf{d}_{3,2} - \mathbf{d}_{3,0}\right) P\left(s_{seg}, t\right) \mathbf{w} dt \\ =& \int_{s_{seg}}^{+\infty} e^{-r(t-s)} \left(\mathbf{d}_{3,2} - \mathbf{d}_{3,0}\right) m \begin{pmatrix} \frac{\partial q_3}{\partial s_0} & 0 & 0 & -\frac{\partial q_3}{\partial s_0} \end{pmatrix} P\left(s_{seg}, t\right) \mathbf{w} dt \\ =& \int_{s_{seg}}^{+\infty} e^{-r(t-s)} m \frac{\partial q_3(t)}{\partial s_0} \left(V_{30} - V_{33}\right) dt > 0,\end{aligned}$$

as $\frac{\partial q_3(t)}{\partial s_0} \leq 0$. For s_{obs},

$$\begin{aligned}&\frac{\partial}{\partial s_0} \int_{s_{obs}}^{+\infty} e^{-r(t-s)} \left(\mathbf{d}_{3,2} - \mathbf{d}_{3,0}\right) P\left(s_{obs}, t\right) \mathbf{w} dt \\ =& \int_{s_{obs}}^{+\infty} e^{-r(t-s)} \left(\mathbf{d}_{3,2} - \mathbf{d}_{3,0}\right) m \frac{\partial}{\partial s_0} \begin{pmatrix} -(q_2 + q_3) & 0 & q_2 & q_3 & 0 & 0 & 0 & 0 \end{pmatrix} P\left(s_{obs}, t\right) \mathbf{w} dt \\ =& \int_{s_{obs}}^{+\infty} e^{-r(t-s)} m \left[\frac{\partial q_3(t)}{\partial s_0} \left(V_{20} - V_{23}\right) + \frac{\partial q_2(t)}{\partial s_0} \left(V_{20} - V_{22}\right)\right] dt > 0\end{aligned}$$

analogously to the proof of lemma 5.

The intuition behind these results is that the later the specialization threshold, the lower the proportion of type 3 workers in the unemployment will be. This makes segregation less desirable for type 3 workers, and reduces the opportunity costs for type 2s of matching with type 1, postponing isolation of the latter.

The rest of the speed effects in the table 3 are proved in the same way.

Later isolation of type 1 workers decreases the proportion of type 2 workers in the unemployment, which makes specialization less attractive.

Later self-segregation of type 3 workers reduces the flow of type 2 workers into retraining, thereby reducing q_3 along the path, and making specialization and isolation less attractive. $\square$

B Efficiency

B.0.1 Socially optimal threshold

Consider a social planner who would be able to dictate agents what kind of training to undertake, and oversee the separation decisions in individual workplace. Then the controls of the planner are the elements of the generator matrix that correspond to transitions from one type to another and to matching choice. Equivalently put, the social planner can choose the specialization threshold as well as other thresholds.

The objective function of the planner is the total welfare. Since in this model utility is transferrable, it is particularly easy to formulate starting in period 0 when all agents are of type 1 and there is full employment:

$$W(0) = \int_0^{+\infty} e^{-rt} \mathbf{d}_{1,1} P(0,t)\, \mathbf{w} dt$$

Suppose we found the socially optimal specialization threshold, s^*. The welfare as computed at that point in time is

$$W(s^*) = \int_{s^*}^{+\infty} e^{-r(t-s^*)} \mathbf{p}(s^*)\, P(s^*,t)\, \mathbf{w} dt,$$

where $P(s^*,t) = \exp\left(\int_{s^*}^{t} G(\tau)\, d\tau\right)$, and G is as in (18).

Consider a variation in social planner's policy at time s^* that postpones the specialization takeoff for a short period h. This variation is summarized by an alternative generator matrix H, which prescribes that everyone who gets a training opportunity choose type 2 in the time interval $[s^*, s^*+h]$.

$$H = \left(\begin{array}{cccc|cccc|cccc}
-(mq_1+mq_2+\rho) & mq_1 & mq_2 & 0 & \rho & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\delta & -\delta & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\delta & 0 & -\delta & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
 & & & & -m & mq_1 & mq_2 & mq_3 & 0 & 0 & 0 & 0 \\
 & \multicolumn{2}{c}{0} & & \delta & -\delta & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & \delta & 0 & -\delta & 0 & 0 & 0 & 0 & 0 \\
 & & & & \delta & 0 & 0 & -\delta & 0 & 0 & 0 & 0 \\
\hline
 & & & & & & & & -m(q_2+q_3) & 0 & mq_2 & mq_3 \\
 & \multicolumn{2}{c}{0} & & & \multicolumn{2}{c}{0} & & 0 & 0 & 0 & 0 \\
 & & & & & & & & \delta & 0 & -\delta & 0 \\
 & & & & & & & & \delta & 0 & 0 & -\delta
\end{array}\right)$$

Welfare under this alternative policy is given by

$$\begin{aligned}\tilde{W}(s^*) &= \int_{s^*}^{s^*+h} e^{-r(t-s^*)}\mathbf{p}(s^*)\exp\left(\int_{s^*}^{s^*+h} H(\tau)\,d\tau\right)\mathbf{w}dt \\ &+ \int_{s^*+h}^{+\infty} e^{-r(t-s^*)}\mathbf{p}(s^*)\exp\left(\int_{s^*}^{s^*+h} H(\tau)\,d\tau + \int_{s^*+h}^{t} G(\tau)\,d\tau\right)\mathbf{w}dt.\end{aligned}$$

Compute the difference between $W(s^*)$ and $\tilde{W}(s^*)$ and divide by h:

$$\frac{\tilde{W}(s^*) - W(s^*)}{h} = \int_{s^*}^{s^*+h} \frac{e^{-r(t-s^*)}\mathbf{p}(s^*)}{h}\left[\exp\left(\int_{s^*}^{s^*+h} H(\tau)\,d\tau\right) - \exp\left(\int_{s^*}^{s^*+h} G(\tau)\,d\tau\right)\right]\mathbf{w}dt$$
$$+\int_{s^*+h}^{+\infty} \frac{e^{-r(t-s^*)}\mathbf{p}(s^*)}{h}\left[\exp\left(\int_{s^*}^{s^*+h} H(\tau)\,d\tau + \int_{s^*+h}^{t} G(\tau)\,d\tau\right) - \exp\left(\int_{s^*}^{t} G(\tau)\,d\tau\right)\right]\mathbf{w}dt$$

For h approaching zero, $\int_{s^*}^{s^*+h} G(\tau)\,d\tau = G(s^*)\,h + o(h)$. Consider the first integral:

$$\begin{aligned}&\int_{s^*}^{s^*+h} \frac{e^{-r(t-s^*)}\mathbf{p}(s^*)}{h}\left[\exp\left(\int_{s^*}^{s^*+h} H(\tau)\,d\tau\right) - \exp\left(\int_{s^*}^{s^*+h} G(\tau)\,d\tau\right)\right]\mathbf{w}dt \\ &= e^{-rh}\mathbf{p}(s^*)\left[I + H(s^*)\,h + o(h) - I - G(s^*)\,h + o(h)\right]\mathbf{w} \\ &= e^{-rh}h\mathbf{p}(s^*)\left[H(s^*) - G(s^*)\right]\mathbf{w} + o(h)\end{aligned}$$

This magnitude is of the order $o(1)$. Consider now the second part of $\left(\tilde{W}(s^*) - W(s^*)\right)/h$:

$$\lim_{h\downarrow 0}\int_{s^*+h}^{+\infty} \frac{e^{-r(t-s^*)}\mathbf{p}(s^*)}{h}\left[\exp\left(\int_{s^*}^{s^*+h} H(\tau)\,d\tau + \int_{s^*+h}^{t} G(\tau)\,d\tau\right) - \exp\left(\int_{s^*}^{t} G(\tau)\,d\tau\right)\right]\mathbf{w}dt$$
$$= \int_{s^*}^{+\infty} e^{-r(t-s^*)}\mathbf{p}(s^*)\left[H(s^*) - G(s^*) + \rho(t-s)\,A\right]\exp\left(\int_{s^*}^{t} G(\tau)\,d\tau\right)\mathbf{w}dt,$$

where

$$A = \left(\begin{array}{cccc|cccc|cccc}
0 & -q_1 & -q_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
 & & & & 0 & -q_1 & -q_2 & q_1+q_2 & 0 & 0 & 0 & 0 \\
 & \multicolumn{2}{c}{0} & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
 & & & & & & & & 0 & 0 & -q_2 & q_1+q_2 \\
 & \multicolumn{2}{c}{0} & & & \multicolumn{2}{c}{0} & & 0 & 0 & 0 & 0 \\
 & & & & & & & & 0 & 0 & 0 & 0 \\
 & & & & & & & & 0 & 0 & 0 & 0
\end{array}\right)$$

Summarizing,

$$\begin{aligned}&\lim_{h\downarrow 0}\frac{\tilde{W}(s^*) - W(s^*)}{h} \\ &= \int_{s^*}^{+\infty} e^{-r(t-s^*)}\mathbf{p}(s^*)\left[H(s^*) - G(s^*) + \rho(t-s)\,A\right]\exp\left(\int_{s^*}^{t} G(\tau)\,d\tau\right)\mathbf{w}dt \qquad (27)\end{aligned}$$

At the command optimum, a small deviation should neither increase nor decrease welfare, so equating (27) to zero gives us the condition on the socially optimal specialization threshold.

B.0.2 Comparison with the decentralized equilibrium

In order to say whether the decentralized equilibrium is efficient, I evaluate (27) at s_0. Consider first the term $\mathbf{p}(s_0)[H(s_0) - G(s_0)]$. It simplifies to

$$\mathbf{p}(s_0)[H(s_0) - G(s_0)] = \rho p_{(1,0)}\mathbf{d}_{2,0} - \rho\left(p_{(1,0)} + p_{(2,0)}\right)\mathbf{d}_{3,0}.$$

Recalling that moment s_0 is defined by $V_{20}(s_0) = V_{30}(s_0)$,

$$\begin{aligned} & \int_{s_0}^{+\infty} e^{-r(t-s_0)}\mathbf{p}(s_0)[H(s_0) - G(s_0)]\exp\left(\int_{s_0}^{t} G(\tau)\,d\tau\right)\mathbf{w}dt \\ = & \ \rho q_1 V_{20} - \rho(q_1 + q_2)V_{30} = -\rho q_2 V_{30} < 0. \end{aligned} \tag{28}$$

Inequality (28) means that in decentralized equilibrium, specialization takes off too late due to the coordination failure among workers.

The second term of (27) simplifies as

$$\begin{aligned} & \rho\int_{s_0}^{+\infty} e^{-r(t-s_0)}\mathbf{p}(s_0)(t-s)A\exp\left(\int_{s_0}^{t} G(\tau)\,d\tau\right)\mathbf{w}dt \\ = & \ \rho U(s_0)\int_{s_0}^{+\infty} e^{-r(t-s_0)}\begin{pmatrix} 0 & -q_1^2 & -q_1q_2\psi_{12} & 0 & 0 & -q_1q_2\psi_{12} & -q_2^2 & q_2 & 0 & 0 & 0 & 0 \end{pmatrix}\mathbf{V}(t)\,dt \\ = & \ \rho U(s_0)\int_{s_0}^{+\infty} e^{-r(t-s_0)}\left(-q_1(q_1V_{11} + q_2V_{12}) - q_2(q_1V_{21} + q_2V_{22}) + q_2V_{23}\right)dt \end{aligned}$$

This expression formalizes another source of inefficiency of decentralized equilibrium. Earlier specialization means more type 3 workers on the market which in turn imply more matches of type $(2,3)$ and fewer matches of of type $(1,1)$, $(1,2)$, and $(2,2)$. Since $y_{11} < y_{12} < y_{22} < y_{23}$, and using the expression for threshold, it follows that the second term is positive.

C Technical lemmata

Lemma 1 *For an arbitrary time-varying row vector* $\mathbf{y}(t)$

$$\int_{s}^{+\infty} e^{-r(t-s)}\left[\int_{s}^{t}\mathbf{y}(\tau)\,dt\right]P(s,t)\,\mathbf{w}dt = \int_{s}^{+\infty} e^{-r(t-s)}\mathbf{y}(t)\,\mathbf{V}(t)\,dt$$

Proof. Define $\boldsymbol{v}(t) = \int_s^t \mathbf{y}(\tau)\,dt$. Then we can integrate by parts:

$$\begin{aligned} I &= \int_{s}^{+\infty} e^{-r(t-s)}\left[\int_{s}^{t}\mathbf{y}(\tau)\,d\tau\right]P(s,t)\,\mathbf{w}dt \\ &= \left(\boldsymbol{v}(\cdot)\int e^{-r(\tau-s)}P(s,\tau)\,\mathbf{w}d\tau\right)\Bigg|_{s}^{\infty} - \int_{s}^{+\infty}\left(\int_{s}^{t} e^{-r(\tau-s)}P(s,\tau)\,\mathbf{w}d\tau\right)d\boldsymbol{v}(t) \end{aligned}$$

Notice that $\int_s^t e^{-r(\tau-s)} P(s,\tau)\, \mathbf{w} d\tau = \mathbf{V}(s) - e^{-r(t-s)}\mathbf{V}(t)$, $d\boldsymbol{\upsilon}(t) = \mathbf{y}(t)\, dt$. So,

$$\begin{aligned} I &= \boldsymbol{\upsilon}(\cdot)\left(\mathbf{V}(s) - e^{-r(t-s)}\mathbf{V}(t)\right)\Big|_s^{\infty} - \int_s^{+\infty} \left[\mathbf{V}(s) - e^{-r(t-s)}\mathbf{V}(t)\right] d\boldsymbol{\upsilon}(t) \\ &= \mathbf{V}(s)\int_s^{\infty} \mathbf{y}(t)\, dt - \int_s^{+\infty} \mathbf{y}(t)\left[\mathbf{V}(s) - e^{-r(t-s)}\mathbf{V}(t)\right] dt \\ &= \int_s^{+\infty} e^{-r(t-s)}\mathbf{y}(t)\, \mathbf{V}(t)\, dt \end{aligned}$$

□

Lemma 2 *Suppose at time 0 unemployment is zero and all workers are of type 1. The dynamics of type 1 unemployment share* before *the specialization takeoff is given by*

$$q_1(t) = \frac{(2m+\delta)\left(e^{\lambda_2 t} - e^{\lambda_1 t}\right)}{\xi\left(1 - e^{-\frac{1}{2}(2m+\delta)t}\right)}$$

for $t \le s_0$ and constant ξ, η, λ_1, λ_2.

Proof. The dynamics of employment and unemployment of type 1 and type 2 workers before the specialization takeoff (under assumption that type 2 workers do not avoid type 1 workers) is governed by the following four differential equations:

$$\begin{pmatrix} \dot{U}_1 \\ \dot{L}_1 \\ \dot{U}_2 \\ \dot{L}_2 \end{pmatrix} = \begin{pmatrix} -\rho - m & \delta/2 & 0 & 0 \\ m & -\delta/2 & 0 & 0 \\ \rho & 0 & -m & \delta/2 \\ 0 & 0 & m & -\delta/2 \end{pmatrix} \begin{pmatrix} U_1 \\ L_1 \\ U_2 \\ L_2 \end{pmatrix}.$$

Solving this system explicitly under initial conditions

$$\begin{cases} U_1(0) = 0 \\ L_1(0) = 1 \\ U_2(0) = 0 \\ L_2(0) = 0 \end{cases}$$

gives

$$\begin{cases} L_1(t) = -\frac{1}{2}\frac{\eta-\xi}{\xi}e^{\lambda_1 t} + \frac{1}{2}\frac{\eta+\xi}{\xi}e^{\lambda_2 t} \\ U_2(t) = \frac{\delta}{2m+\delta} + \delta\frac{e^{\lambda_1 t} - e^{\lambda_2 t}}{\xi} - \frac{\delta}{2m+\delta}e^{-\frac{1}{2}(2m+\delta)t} \\ U_1(t) = \delta\frac{e^{\lambda_1 t} - e^{\lambda_2 t}}{\xi} \\ L_2(t) = \frac{2m}{2m+\delta} + \frac{1}{2}\frac{\eta-\xi}{\xi}e^{\lambda_1 t} - \frac{1}{2}\frac{\eta+\xi}{\xi}e^{\lambda_2 t} + \frac{\delta}{2m+\delta}e^{-\frac{1}{2}(2m+\delta)t} \end{cases}$$

where

$$\begin{aligned} &\eta = 2\rho + 2m - \delta, \\ &\xi = \sqrt{\eta^2 + 8\delta m}, \\ &\lambda_1 = -\tfrac{2\delta+\eta+\xi}{4}, \\ &\lambda_2 = -\tfrac{2\delta+\eta-\xi}{4} > \lambda_1. \end{aligned}$$

This implies

$$q_1(t) = \frac{U_1}{U_1 + U_2} = \frac{(2m+\delta)\left(e^{\lambda_2 t} - e^{\lambda_1 t}\right)}{\xi\left(1 - e^{-\frac{1}{2}(2m+\delta)t}\right)}. \tag{29}$$

□

Lemma 3 *For all i and $j > 0$ the following holds at all times:*

$$V_{ij}(s) = \int_s^{+\infty} \left[\int_s^t w_{ij} e^{-(\tau - s)r} d\tau + V_{i0} e^{-(t-s)r}\right] dF(t), \tag{30}$$

where $F(t)$ is the cdf of exponential distribution with parameter δ.

Proof. Worker of type i while matched with type j receives flow benefit w_{ij} until their match dissolves, whereupon her/his discounted expected utility is given by V_{i0}. (30) follows immediately from the fact that waiting time before match dissolution is distributed exponentially with parameter δ. □

Lemma 4 *Consider row vector* $\mathbf{y} = m\begin{pmatrix} 0 & \frac{\partial q_1}{\partial x} & \frac{\partial q_2}{\partial x} & \frac{\partial q_3}{\partial x} & -\frac{\partial q_1}{\partial x} & 0 & -\frac{\partial q_2}{\partial x} & -\frac{\partial q_3}{\partial x} \end{pmatrix}$, *for arbitrary parameter x. Then,*

1. *If $\frac{\partial q_1}{\partial x} < 0$, $\frac{\partial q_2}{\partial x} < 0$, and $\frac{\partial q_3}{\partial x} > 0$,*

$$\int_{s_0}^{+\infty} e^{-r(t-s)} \mathbf{y}(t)\, \mathbf{V}(t)\, dt < 0.$$

2. *If $\frac{\partial q_1}{\partial x} > 0$, $\frac{\partial q_2}{\partial x} > 0$, and $\frac{\partial q_3}{\partial x} < 0$,*

$$\int_{s_0}^{+\infty} e^{-r(t-s)} \mathbf{y}(t)\, \mathbf{V}(t)\, dt > 0.$$

Proof. Consider the product

$$\begin{aligned} \mathbf{y}(t)\, \mathbf{V}(t) &= m\left(\frac{\partial q_1}{\partial x} V_{21} + \frac{\partial q_2}{\partial x} V_{22} + \frac{\partial q_3}{\partial x} V_{23} - \frac{\partial q_1}{\partial x} V_{30} - \frac{\partial q_2}{\partial x} V_{32} - \frac{\partial q_3}{\partial x} V_{33}\right) \\ &= m\left(-\frac{\partial q_3}{\partial x}\left[(V_{33} - V_{30}) - (V_{23} - V_{21})\right] - \frac{\partial q_2}{\partial x}\left[(V_{32} - V_{30}) - (V_{22} - V_{21})\right]\right) \end{aligned} \tag{31}$$

The last equality was derived using the identity $\frac{\partial q_1}{\partial x} + \frac{\partial q_2}{\partial x} + \frac{\partial q_3}{\partial x} = 0$ (Since $q_1 + q_2 + q_3 \equiv 1$). Let us demonstrate that (31) is negative. By lemma 3,

$$\begin{aligned} & V_{33} - V_{32} - (V_{23} - V_{22}) \\ = & \int_s^{+\infty} \left[\int (w_{33} - w_{32} - w_{23} + w_{22})\, e^{-(\tau - s)r} d\tau + (V_{30} - V_{20})\, e^{-(t-s)r}\right] dF(t) \end{aligned} \tag{32}$$

At and after the specialization threshold $V_{30}(t) - V_{20}(t) \geq 0$; $w_{33} - w_{32} - w_{23} + w_{22} = (y_{33} - 2y_{23} + y_{22})/2 > 0$ by assumption (1). Thus, (32) is positive and

$$\begin{aligned}
& V_{33} - V_{32} - (V_{23} - V_{22}) > 0 \\
\Rightarrow\ & (V_{33} - V_{30}) - (V_{23} - V_{21}) > (V_{32} - V_{30}) - (V_{22} - V_{21}) \\
\Rightarrow\ & \frac{\partial q_3}{\partial x}\left[(V_{33} - V_{30}) - (V_{23} - V_{21})\right] > -\frac{\partial q_2}{\partial x}\left[(V_{32} - V_{30}) - (V_{22} - V_{21})\right] \\
\Leftrightarrow\ & -\frac{\partial q_3}{\partial x}\left[(V_{33} - V_{30}) - (V_{23} - V_{21})\right] - \frac{\partial q_2}{\partial x}\left[(V_{32} - V_{30}) - (V_{22} - V_{21})\right] < 0
\end{aligned}$$

The third inequality was obtained by observing that $\frac{\partial q_3}{\partial x} > -\frac{\partial q_2}{\partial x} > 0$. Hence, for all $t \geq s_0$

$$\begin{aligned}
& \mathbf{y}(\mathbf{t})\,\mathbf{V}(\mathbf{t}) < 0 \\
\Rightarrow\ & \int_{s_0}^{+\infty} e^{-r(t-s)}\mathbf{y}(t)\,\mathbf{V}(\mathbf{t})dt < 0.
\end{aligned}$$

The proof of part 2) proceeds in the same way, observing that $\frac{\partial q_3}{\partial x} < -\frac{\partial q_2}{\partial x} < 0$.□

Lemma 5 *Starting from the specialization threshold, for $t > s_0$,*

$$\begin{aligned}
& \frac{\partial q_3(t)}{\partial \rho} > 0, \frac{\partial q_1(t)}{\partial \rho} < 0, \frac{\partial q_2(t)}{\partial \rho} < 0; \\
& \frac{\partial q_3(t)}{\partial \delta} < 0, \frac{\partial q_1(t)}{\partial \delta} > 0, \frac{\partial q_2(t)}{\partial \delta} > 0; \\
\textit{if } q_1 > \frac{1}{2},\ & \frac{\partial q_3(t)}{\partial m} > 0, \frac{\partial q_1(t)}{\partial m} < 0, \frac{\partial q_2(t)}{\partial m} < 0.
\end{aligned}$$

Proof. After the threshold, the dynamics of the pools of employed and unemployed workers of type 1, 2, 3 is governed by the following six equations:

$$\begin{aligned}
\dot{U}_1 &= -\rho U_1 - m\left(\tfrac{U_1 + U_2\psi_{12}}{U_1+U_2+U_3}\right)U_1 + L_1\delta/2 \\
\dot{L}_1 &= m\left(\tfrac{U_1 + U_2\psi_{12}}{U_1+U_2+U_3}\right)U_1 - L_1\delta/2 \\
\dot{U}_2 &= -\rho U_2 - m\left(\tfrac{U_1\psi_{12} + U_2 + U_3\psi_{23}}{U_1+U_2+U_3}\right)U_2 + L_2\delta/2 \\
\dot{L}_2 &= m\left(\tfrac{U_1\psi_{12} + U_2 + U_3\psi_{23}}{U_1+U_2+U_3}\right)U_2 - L_2\delta/2 \\
\dot{U}_3 &= \rho U_1 + \rho U_2 - m\left(\tfrac{U_3 + U_2\psi_{23}}{U_1+U_2+U_3}\right)U_3 + L_3\delta/2 \\
\dot{L}_3 &= m\left(\tfrac{U_3 + U_2\psi_{23}}{U_1+U_2+U_3}\right)U_3 - L_3\delta/2
\end{aligned}$$

In the case when $\psi_{12} = \psi_{23} = 1$, we can rewrite this system in terms of proportions q_1, q_2, q_3, noting that $U = 1 - L_1 - L_2 - L_3$ (constant population):

$$\begin{aligned}
\dot{q}_1 &= -\rho q_1 + q_1 m\left(-q_1 + q_1^2 + q_1 q_2 + q_3^2 + q_3 q_2\right) + \tfrac{\delta}{2U}\left((1-q_1)L_1 - q_1 L_2 - q_1 L_3\right) \\
\dot{L}_1 &= m\left(q_1^2 + q_1 q_2\right)U - L_1\delta/2 \\
\dot{q}_2 &= -\rho q_2 + q_2 m\left(-1 + q_1^2 + q_1 q_2 + q_2 + q_3^2 + q_3 q_2\right) + \tfrac{\delta}{2U}\left((1-q_2)L_2 - q_2 L_1 - q_2 L_3\right) \\
\dot{L}_2 &= m q_2 U - L_2\delta/2 \\
\dot{q}_3 &= \rho(q_1 + q_2) + q_3 m\left(q_1^2 + q_1 q_2 + q_3^2 + q_3 q_2 - q_3\right) + \tfrac{\delta}{2U}\left((1-q_3)L_3 - q_3 L_1 - q_3 L_2\right) \\
\dot{L}_3 &= m\left(q_2 q_3 + q_3^2\right)U - L_3\delta/2
\end{aligned} \quad (33)$$

The initial conditions are $q_1(s_0)$, $q_2(s_0)$, $q_3(s_0) = 0$ and similarly for $L_i(s_0)$. This system of differential equations in non-linear and unsolvable in closed form. But we are only interested in signs of the derivatives $\frac{\partial q_j(t)}{\partial x}$.

Consider a system of differential equations

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \boldsymbol{\mu}), \tag{34}$$

with initial conditions $\mathbf{x}(0) = \mathbf{x}_0$, where $\mathbf{x}$ is a vector of dimension n and $\boldsymbol{\mu}$ is a vector of parameters.

If $\mathbf{f}(\cdot)$ is differentiable, then derivatives of the solution $\mathbf{x}(t)$ to (34) w.r.t. parameter μ_k exist and are continuous [Pontryagin, 1965]. Furthermore, these derivatives $\frac{\partial x^i(t)}{\partial \mu_k} = \psi_k^i$ are themselves solutions to the differential equations

$$\dot{\psi}_i^k = \sum_j \frac{\partial f^i}{\partial x_j}\psi^j + \frac{\partial f^i}{\partial \mu_k},\ i = 1, .., n \tag{35}$$

with the initial conditions $\psi_i^k(0) = 0\ \forall i$. System (35) is called the associated differential system in variations.

Let us set up the system in variations associated with (33) for parameter ρ:

$\dot{\psi}_1 = \sum_j \frac{\partial f^1}{\partial x_j}\psi_j - q_1$

$\dot{\psi}_2 = \sum_j \frac{\partial f^2}{\partial x_j}\psi_j - q_2$

$\dot{\psi}_3 = \sum_j \frac{\partial f^3}{\partial x_j}\psi_j + q_1 + q_2$

Given the initial conditions $\psi_i(s_0) = 0$, this implies $\frac{\partial q_1}{\partial \rho} = \psi_1 < 0$, $\frac{\partial q_2}{\partial \rho} = \psi_2 < 0$, $\frac{\partial q_3}{\partial \rho} = \psi_3 < 0$ for $t > s_0$.

The proofs for the derivatives w.r.t. δ and m are obtained analogously. The associated systems in variations are given by

with respect to δ: $\begin{cases} \dot{\psi}_1 = \sum_j \frac{\partial f^1}{\partial x_j}\psi_j + \frac{1}{2U}\left((1-q_1)L_1 - q_1L_2 - q_1L_3\right) \\ \dot{\psi}_2 = \sum_j \frac{\partial f^2}{\partial x_j}\psi_j + \frac{1}{2U}\left((1-q_2)L_2 - q_2L_1 - q_2L_3\right) \\ \dot{\psi}_3 = \sum_j \frac{\partial f^3}{\partial x_j}\psi_j + \frac{1}{2U}\left((1-q_3)L_3 - q_3L_1 - q_3L_2\right) \end{cases}$

Thus, $\frac{\partial q_3(t)}{\partial \delta} < 0$, $\frac{\partial q_1(t)}{\partial \delta} > 0$, $\frac{\partial q_2(t)}{\partial \delta} > 0$.

With respect to m: $\begin{cases} \dot{\psi}_1 = \sum_j \frac{\partial f^1}{\partial x_j}\psi_j + q_1q_3(1-2q_1), \\ \dot{\psi}_2 = \sum_j \frac{\partial f^2}{\partial x_j}\psi_j - 2q_1q_2q_3, \\ \dot{\psi}_3 = \sum_j \frac{\partial f^3}{\partial x_j}\psi_j + q_3q_1(1-2q_3), \end{cases}$

For $q_1 > \frac{1}{2}$, $\frac{\partial q_3(t)}{\partial m} > 0$, $\frac{\partial q_1(t)}{\partial m} < 0$, $\frac{\partial q_2(t)}{\partial m} < 0$. $\square$

References

Acemoglu, D. [1996a]. Changes in unemployment and wage inequality: An alternative theory and some evidence?, *Technical Report 1459*, Centre for Economic Policy Research. Discussion Paper #1459.

Acemoglu, D. [1996b]. A microfoundation for social increasing returns in human capital accumulation, *Quarterly Journal of Economics* **111**: 779–804.

Acemoglu, D. [1997]. Training and innovation in an imperfect labor market, *Review of Economic Studies* **64**: 445–464.

Acemoglu, D. and Zilibotti, F. [1998]. TFP differences. MIT mimeo.

Aghion, P. and Blanchard, O. J. [1994]. On the speed of transition in Central Europe, *NBER Macroeconomics Annual*, Cambridge, MA.

Bewley, T. [1995]. A depressed labor market as explained by participants, *American Economic Review* **85**: 250–254.

Bewley, T. [1998]. Why not cut pay?, *European Economic Review* **42**: 459–90.

Blanchard, O. J. [1997]. *The Economics of Post-Communist Transition*, Clarendon Press, Oxford.

Blanchard, O. and Kremer, M. [1997]. Disorganization, *Quarterly Journal of Economics* .

Boeri, T. [1999]. Transition with labor supply. Paper presented at the 5th Nobel Symposium in Economics.

Boeri, T., Burda, M. C. and Köllõ, J. [1998]. *Mediating the Transition: Labour Markets in Central and Eastern Europe*, Institute for EastWest Studies.

Bordin, A. [1998]. Labour Market in Today's Russia: Its History and Geography, *The New Generation* **2**.

Brixiova, Z. and Li, W. [1998]. Skill acquisition and private firm creation in transition economies, William Davidson Institute, Working paper #162.

Burdett, K. and Coles, M. [1997]. Marriage and class, *Quarterly Journal of Economics* **112**: 141–168.

Caballero, R. and Hammour, M. [1996]. On the ills of adjustment, *Journal of Development Economics* **51**(1).

Castanheira, M. and Roland, G. [1997]. The optimal speed of transition: A general equilibrium analysis. presented at the first annual EERC conference "Economic Transition in CEE and Russia, Theory and Empirical Evidence".

Cazes, S. and Scarpetta, S. [1998]. Labour Market Transitions and Unemployment Duration: Evidence from Bulgarian and Polish Micro-Data, *Economics of Transition* **6**: 113–144.

Coricelli, F., Hegeemejer, K. and Rybinski, K. [1995]. Poland, *in* S. Commander and F. Coricelli (eds), *Unemployment, Restructuring and the Labor Market in Eastern Europe and Russia*, EDI Development Studies. The World Bank, chapter 2.

Fan, C. S., Overland, J. and Spagat, M. [1997]. Human capital, growth, and inequality in Russia. University of London, mimeo.

Flanagan, R. J. [1998]. Were communists good human capitalists? The case of the Czech Republic, *Labor Economics* **5**: 295–312.

Ham, J., Svejnar, J. and Terrell, K. [1995]. Czech Republic and Slovakia, *in* S. Commander and F. Coricelli (eds), *Unemployment, Restructuring and the Labor Market in Eastern Europe and Russia*, EDI Development Studies. The World Bank, chapter 3, pp. 91–146.

Hunt, J. [1999]. Post-unification wage growth in East Germany. NBER Working Paper No. 6878.

Kim, S. [1989]. Labor specialization and the extent of the market, *Journal of Political Economy* **97**(3).

Kremer, M. [1993]. The O-ring theory of economic development, *Quarterly Journal of Economics* **108**: 551–75.

Kremer, M. and Maskin, E. [1996]. Wage inequality and segregation by skill, National Bureau of Economic Research, Working paper # 5718.

Kremer, M. and Thomson, J. [1998]. Why isn't convergence instantaneous? Young workers, old workers, and gradual adjustment, *Journal of Economic Growth* **3**: 5–28.

Krueger, A. B. and Pischke, J. S. [1995]. A comparative analysis of East and West German labor markets: Before and after unification, *in* R. B. Freeman and L. F. Katz (eds), *Differences and Changes in Wage Structures*, The University of Chicago Press, chapter 12, pp. 405–445.

Kuznets, S. [1955]. Economic growth and income inequality, *American Economic Review* **XLV**: 1–28.

Lu, X. and McAfee, R. P. [1996]. Matching and expectations in a market with heterogeneous agents, *in* M. R. Baye (ed.), *Advances in Applied Microeconomics*, Vol. 6, Greenwich, pp. 121–156.

Münich, D., Svejnar, J. and Terrell, K. [1998]. Worker-firm matching and unemployment in transition to a market economy: (Why) are the Czechs more successful than others? CERGE-EI mimeo.

Noorkoiv, Orazem, Puur and Vodopivec [1998]. Employment and wage dynamics in Estonia, 1989-95, *Economics of Transition* **6**(2): 481–503.

Orazem, P. and Vodopivec, M. [1997]. Value of human capital in transition to market: Evidence from Slovenia, *European Economic Review* **41**(3-5).

Pissarides, C. [1990]. *Equilibrium Unemployment Theory*, Basil Blackwell.

Pontryagin, L. S. [1965]. *A Course in Ordinary Differential Equations*, Hindustan Pub. Corp., Delhi.

Rutkowski, J. [1996]. High skills pay off: The changing wage structure during economic transition in Poland, *Economics of Transition* **4**(1).

Rutkowski, J. [1998a]. Changes in the wage structure during economic transition in Central and Eastern Europe. World Bank technical paper #340.

Rutkowski, J. [1998b]. Welfare and the labor market in Poland. World Bank technical Paper #417.

Saint-Paul, G. [1996]. *Dual Labor Markets*, MIT Press, Cambridge, MA.

Sarychev, A. [1999]. Analytical notes on the adjustment of human capital in a setting with profit-maximizaing firms. MIT mimeo.

Shimer, R. and Smith, L. [2000]. Assortative matching and search, *Econometrica* **forthcoming**.

The World Bank Development Report [1996]. *From Plan to Market*, Oxford University Press.

Topkis, D. [1998]. *Supermodularity and Complementarity*, Princeton University Press.

Večerník, J. [1995]. Changing earnings distribution in the Czech Republic: Survey evidence from 1988-1994, *Economics of Transition* **3**: 355–71.