QUT School of Economics and Finance

Discussion Paper and Working Paper Series

Unemployment benefits and educational choices

Paul Frijters

Abstract:

In this paper it is argued that the risk-pooling role of unemployment benefits affects the irreversible choices of future labour market entrants. One such ex ante choice is education. Some types of education lead to general human capital that lead to almost certain employment. Other types of education are more specialised and lead to less secure employment. We address this issue in OLG search models that allow for risk-aversion, heterogeneity in talents, endogenous price formation of different specialisations, and competitive wage formation. We find that in the absence of unemployment benefits, the percentage of individuals taking high-risk specialised education is inefficiently low. Those with higher innate abilities are typically found to take lower degrees of specialisation, implying that the relation between wages and risks at the individual level is the reverse from what it is at the aggregate level. We find that an unemployment benefit (UB) system raises efficiency and welfare because it promotes efficient specialisation. Because education takes time, it takes a long time before the composition of the workforce has adapted to changing incentives. With a calibrated model we explore such lags between unexpected changes in circumstances and outcomes.

Unemployment benefits and educational choices

Paul Frijters[*]
Preliminary. Do not quote.

November 30, 2001

Abstract

In this paper it is argued that the risk-pooling role of unemployment benefits affects the irreversible choices of future labour market entrants. One such ex ante choice is education. Some types of education lead to general human capital that lead to almost certain employment. Other types of education are more specialised and lead to less secure employment. We address this issue in OLG search models that allow for risk-aversion, heterogeneity in talents, endogenous price formation of different specialisations, and competitive wage formation. We find that in the absence of unemployment benefits, the percentage of individuals taking high-risk specialised education is inefficiently low. Those with higher innate abilities are typically found to take lower degrees of specialisation, implying that the relation between wages and risks at the individual level is the reverse from what it is at the aggregate level. We find that an unemployment benefit (UB) system raises efficiency and welfare because it promotes efficient specialisation. Because education takes time, it takes a long time before the composition of the workforce has adapted to changing incentives. With a calibrated model we explore such lags between unexpected changes in circumstances and outcomes.

1 Introduction and literature

Many studies have looked at the possible advantages of unemployment benefits (UB) on the decisions or welfare of an individual after he has entered the labour market. One such advantage, empirically assessed by Gruber (1997), is that UB smooths incomes. Another potential benefit is that UB increases the outside option of unemployed individuals and hence allows them to make better decisions (Mortensen and Pissarides, 1999). Acemoglu and Shimer (1999, 2000) show that UB give an added incentive for unemployed individuals to hold out for jobs that involve a higher risk of lay-off or that are simply harder to find. In their model, this can raise welfare and output for low levels of UB. In a similar vein Marimon and Zilibotti (1999), extending the basic argument made at least as early as Burdett (1979), show that when individuals receive unemployment benefits while unemployed, they have a greater incentive to hold out for 'better matches', i.e., jobs that are higher paying. The fact that holding out for better matches has an external beneficial effect on the probability that others find their optimal match, can make UB welfare and output increasing. It is a disputed empirical question when and whether such advantages of UB are enough to outweigh the decreased incentives to search jobs in the presence of moral hazard (e.g. Hopenhayn and Nicolini, 1997).

In this paper, we add to this literature by looking at the effects that unemployment benefits have on the economy via its effect on irreversible choices made before entering the labour market, such as education. Education not only affects the future level of expected wages, but also ties an individual to the unemployment risks of the jobs that an individual is then suited for. Individuals who opt for general education can apply to and function in many jobs. Hence, their job-arrival rate will be high once they enter the labour market and their job destruction rate low. Individuals who opt for specialist skills however will have a much smaller pool of jobs to search and are more vulnerable to technology shocks that wipe out the value of specialist knowledge. Hence the choice for type of education is also a choice for a level of future risk. This choice may therefore change if the risks or rewards on the labour market change. The presence of a risk-sharing device like unemployment benefits reduces (the importance of) these risks and therefore will affect educational choice. This in turn will change unemployment dynamics and average outcomes, such as average output, and relative wages.

We analyse the importance of risk-relevant irreversible early choices in a

general equilibrium OLG-model. In the simplest version, identical risk-averse individuals who live only two periods choose a degree of specialisation in the first period. Greater specialisation translates to an increased risk of unemployment in the second period. Those who work with a specific specialisation make intermediaries that are transformed into a homogeneous consumption good by a continuous CES production function. This production functions brings out macro-economic complementarities between the activities of individuals that so far have always been ignored in search models. The model yields an observed wage distribution that is unique and that can have virtually any shape, depending on the underlying functions.

Introducing unemployment benefits paid for by a single marginal tax rate turns out to make investments into more risky professions relatively more attractive. This tilts the distribution of specialisation choices to the right, which increases production, increases expected utility, and increases unemployment. In this regard, the insurance provided by UB has the same effect in this model as in the models by Acemoglu and Shimer (1999, 2000) and Marimon and Zillibotti (1999), albeit via a different mechanism. An empirical prediction of these findings is that the percentage of individuals taking general education is less in an area with higher unemployment benefits. If we view the US as an area with relatively low unemployment benefits and the countries in the EU as an area with high unemployment benefits, this would translate in there being more vocational and specialist education in the EU than in the US. Ashton and Green (1996) have indeed documented such a diference.

Apart from risks and macro-economic complementarities, innate talents are also an important determinant of wages. The two-period model is therefore generalised to include talent heterogeneity, where talent is understood as an efficiency unit of productivity for any specialisation. This does not alter the previous conclusions bar one: it then turns out that under most utility functions, those with highest total wages are then also the ones with least risks. Those with most talent 'buy' higher security by taking low-risk jobs whilst their total wages are still higher because of their higher talent levels. This extended model is hence capable of reconciling the theoretical prediction that individuals would have to get higher wages when running higher risks and the empirical evidence that suggests that those with higher risks actually get paid less (e.g. Hwang et al. 1992).

To study the dynamics of the model, we first extend the two-period model to an infinite period continuous-time OLG model where a higher degree of

specialisation implies lower job-finding rates and higher job-destruction rates. For the most plausible parametric assumptions, the main conclusions from the two-period models apply to this continuous case also.

We then calibrate a dynamic 20-period OLG model in section 3, with which we can study the dynamics of the economy after unanticipated shocks. We find that, because educational choices are irreversible, changes in parameters tend to have delayed effects. The effect of reducing taxes and unemployment benefits on unemployment takes much longer to work through when there is a large long-run decrease in the number of specialists, simply because it takes several cohorts before the new equilibrium mix between generalists and specialists is achieved. This explanation of unemployment rates combines shocks with institutional factors, is in line with the recommendations of Blanchard and Wolfers (2001).

General economic shocks turn out to have a longer lasting effect (persistence) in the economy with high unemployment benefits than in the economy without. This is because the economy with unemployment benefits has more specialists who have lower job-finding rates. Hence, after an economic shock it takes longer for these individuals to find a job again than it takes the generalists, of which there are more in the economy with low unemployment benefits. This is another explanation for the sluggishness with which the EU unemployment rates came down after the oil price shocks in the 70's (Blanchard en Wolfers, 1999). Sargent and Ljungqvist (1998) explained this same phenomenon by arguing that unemployment benefits increased the willingness of individuals to wait for better jobs. Sargent and Ljungqvist combined this with the assumption that individual loose skills in unemployment. Hence unemployment benefits aggravated the negative shocks in their model because the high unemployment benefits indirectly adversely changed the characteristics of individuals whilst they were already on the labour market. The difference with this paper is that in this paper, there is no resort to a reduction in skills during unemployment, but an effect of unemployment benefits on the long term composition of the workforce. Hence this paper stresses that the US and EU labour market were already very different before the oil price shocks, even if at that moment unemployment benefits would have been equated. The evidence for the existence of loss-of-skills or stigma as a result of prolonged unemployment is still very scant (see e.g. Heckman and Borjas, 1980, Lynch; Frijters, an den Berg and Lindeboom 2001, Bonnal et al. 1997). This favours explanations of persistent unemployment that do not require changes in characteristics of individuals during unemployment.

The final section concludes.

This paper's theoretical contribution is in three main directions. We believe we are the first to allow for macro-economic production complementarities in a search model. Previous search models have taken the productivity of a specific match to be independent of the activities of individuals in different types of activities (Burdett and Mortensen 1998; Moens 1997; Acemoglu and Shimer 1998; Marimon and Zillibotti 1999; Pissarides 1990 and extensions, such as Fredriksson and Holmlund 2001 or Petrongolo 2001). Second, we solve for entire wage distributions in the context of individual heterogeneity and competitive wage setting. This sets us apart from the extensions of the Burdett-Mortensen (1998) model that also solve for wage distributions and individual heterogeneity (e.g. Bontemps et al 2001) but assume that labour markets are segmented and that not all firms are at the production frontier. It also sets us apart from Acemoglu and Shimer (1999) and Moens (1997) who also have competitive wage setting, but do not solve for distributions and have no heterogeneity in talents. Finally, this paper is the first to solve a continuous time OLG model with risk-aversion, heterogeneity, arbitrary utility functions, job search and job destruction.

2 A 2-period general equilibrium OLG model

There is a continuum of individuals each period of measure 1. A generation of ex ante identical individuals lives two periods, where one generation is in period 1 whilst the previous generation is in period 2. In the first period, individuals can choose a level of specialisation $1 \geq \theta \geq 0$ as their type of skill in the second period. The cumulative density of individuals at time t choosing specialisation θ is denoted by $F_t(\theta)$ where $\int_0^1 dF_t(\theta) = 1$. Looking only at steady states, we will drop the time subscripts.

In the second period, individuals search production facilities. The probability of finding and keeping a job is $h(\theta)$ where $h(0) = 1$, $h(1) > 0$ and $\frac{\partial h}{\partial \theta} < 0$. The specialised jobs are hence by definition harder to find and\or keep. A worker produces one unit of a specialisation-specific good which is an intermediary into a final good, where the production technology is CES. Total production is

$$y = [\int_0^1 (f(\theta)h(\theta))^\gamma d\theta]^{1/\gamma} \tag{1}$$

where $f(\theta)h(\theta)$ is the total amount of the intermediary good of type θ that is produced and where $1 > \gamma > 0$.

Individuals are forward looking risk-averse rational utility maximizers. Expected utility is

$$E\{U(\theta)\} = h(\theta)u((1 - \tau)w(\theta)) + (1 - h(\theta))u(b) \qquad (2)$$

where τ is the tax rate, and b=$\frac{\int \tau y}{(1-h)fd\theta}$ is the level of unemployment benefits. Wages are set competitively: $w(\theta) = \frac{\partial y}{\partial f(\theta)h(\theta)} = (f(\theta)h(\theta))^{\gamma-1}y^{1-\gamma}$ when $f(\theta)$ exists and $w(\theta) = 0$ at positive mass points of $F(\theta)$. Also, $u(.)$ is convex, its derivative exists and is continuous, with $u(\infty) = \infty$. Non-negative non-work incomes ensure that u(0)>-∞.[1] We restrict initial attention to cases where $b < w(\theta)$.

First, a standard argument holds that in equilibrium, $f(\theta)$ has to be continuous if $h(\theta)$ is continuous. The reason is that if $f(\theta)$ is not continuous and, say, drops at some point x, then wages will make a discontinuous jump at x. There is then a first order gain to be made for individuals just before x to change their choice of θ to x, with only a second-order loss of finding a job.

It is also the case that f(θ)>0 for every θ and every utility function because $w(\theta) = \infty$ when f(θ)=0. Choosing θ will hence be preferred over choices with wages less than infinite, which ensures that f(θ)>0 for any θ. These regularities in $f(\theta)$ mean that $w(\theta)$ is also continuous when $h(\theta)$ is continuous.

Having hence checked that minimal regularity conditions apply, we can now characterise the equilibrium by noting that each choice of θ must yield the same expected utility. Differentiating $E\{U(\theta)\}$ with respect to θ and setting to 0 gives the main solution equation of this model:

$$-h'(\theta)(u((1 - \tau)w(\theta)) - u(b)) = (1 - \tau)w'(\theta)h(\theta)u'((1 - \tau)w(\theta)) \qquad (3)$$

for differentiable points of $h(\theta)$. This equation immediately shows that wages are increasing in the risk ($= 1 - h$) and hence increasing in θ. In order to judge the efficiency of the outcome, consider what would be the output maximizing choice of $f(\theta)$, denoted as $f^e(\theta)$. Due to the properties of the

---

[1]This gives a lower bound to the utility function. Without a bound, we would run into the St. Petersburg paradox, where individuals would not have complete preferences. For a short discussion on this problem with Von Neumann - Morgenstern expected utility functions, which was first noted by Savage, see Aumann (1977).

CES-function, y is maximised when $h(\theta)w(\theta)$ is constant. Hence $w^e(\theta) \propto \frac{1}{h(\theta)}$ and $\frac{dw^e}{d(1-h)} = \frac{w^e}{h}$. This corresponds to $f^e(\theta) \propto h(\theta)^{\frac{-\gamma}{\gamma-1}}$. The equilibrium of the model is now characterised in proposition 1.

Proposition 1. *At $\tau = 0$, the model above has a unique equilibrium solution $f(\theta)$, which is inefficient. Any continuous distribution of observed wages $z(w)$ can be supported as long as $z(x) = 0$ for all $0 < x < w_{\mathsf{min}}$.*

Proof: existence and uniqueness is proven in the Appendix. Inefficiency can be seen by noting that we can differentiate $E\{U(\theta)\}$ with respect to h, obtaining $\frac{dw}{d(1-h)} = \frac{u(w(\theta))-u(0)}{h(\theta)u'(w(h))} > \frac{w}{h}$ because of the risk-aversion in $u(.)$. This in turn implies inefficiency of the equilibrium. The shape of observed wages $z(w)$ follows because we can write $z(x) = \frac{1}{\frac{\partial h(\theta)}{\partial \theta}}f(\theta)$ where $\theta = \arg_\theta w(\theta) = x$. Because nothing bounds $\frac{w(1)}{w(0)}$ from above, any observed continuous density function that is bounded from below can then be supported by an appropriate choice of $h(\theta)$ and $u(.)$.

The intuition behind existence is that the main solution equation uniquely maps $w(\theta)$ as a continuous function of $w(0)$. Conversely, this leads to a unique $f(\theta)$ and y, both continuous in $w(0)$. Because $w(0)$ is itself uniquely determined by $f(0)$ and y, there is a closing equation for which a fixed point argument shows it has a solution for at least one $w(0)$. Uniqueness then follows because it is not possible to change $f(\theta)$ without increasing some wages and decreasing others (proven in Lemma 1 in the Appendix). Because equilibrium implies that individuals have equal utility, it cannot be the case that there is a second equilibrium in which there are some better off and others strictly worse off.

The question now is whether this outcome can be improved upon by introducing an unemployment benefit. Three general results can be obtained:

Proposition 2. *i) At 0, an increase in τ is both utility, unemployment and output increasing. ii) There is a critical level τ^z above which all individuals would prefer not to work where τ^z solves: $(1 - \tau)w(0) = b$ and where $w(\theta) = w(0)$ and $f(\theta) \propto h(\theta)^{-1}$. iii) There is no level of τ that yields efficiency.*

7

Proof of iii) Suppose there is a τ that maximizes output it would have to mean that $\frac{dw}{d(1-h)} = \frac{u((1-\tau)w(h))-u(b)}{(1-\tau)hu'((1-\tau)w(h))} = \frac{w}{h}$ for any θ. In turn this would imply that $u(x) - u(b) = xu'(x)$ where $x = (1-\tau)w(\theta)$. This equation in turn can only hold for a continuum of x when $u''(.) = 0$ and $\tau = 0$, which implies that individuals would have to be risk-neutral which contradicts the primitives of the model. The proof of i) and ii) is in the Appendix.

The intuition of this result is that at low levels of tax, the introduction of a benefit induces individuals to take more risks, which increases output, increases unemployment, and increases utility. Because the utility of each choice is the same, the utility increase following an increase in τ is ex ante the same for each individual. Because of the irreversibility of specialisation however, ex post some individuals will not want taxation. The individuals with $\theta = 0$ for instance run run no risks ex post and will hence oppose any tax ex post, even though they have benefitted from it ex ante.

The reason that there is no tax level that will yield the maximum output is that ex post individuals want different levels of insurance: given that individuals will choose different risks, the optimal insurance should differ for different levels of risk. The ex post risk-pooling between individuals with different risks means that those who run little risks will be over insured and those who run high risks will be under insured. This shows up a basic difference between considering wage distributions in stead of a single wage outcome, such as Acemoglu and Shimer (1999), where there is a single tax that restores efficiency.

The key features of this static model are brought out in Figures 1, 2 and 3. Figure 1 shows the relation between $f(\theta)$ and τ.
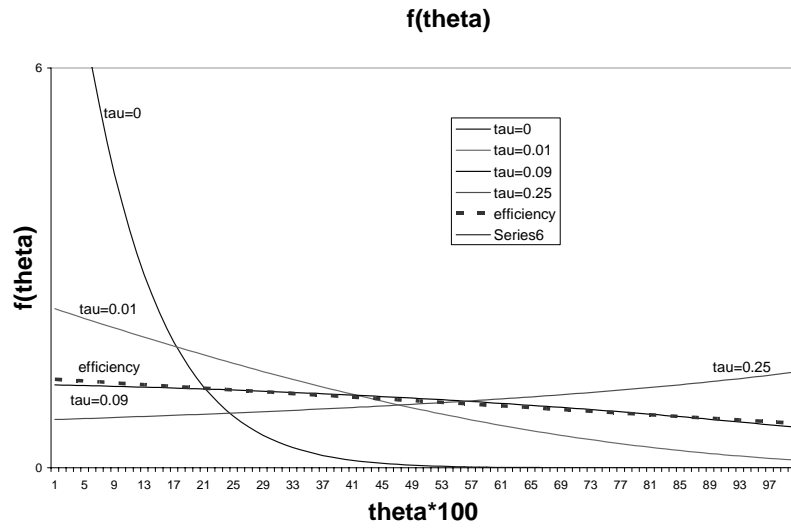
8

f(theta)

Figure 1: the relation between risks and the distribution of risk choices. Parametric assumptions: $h(\theta) = 1 - 0.5 * \theta$, $\gamma = 0.5$, and $u(y) = y^{0.1}$

The straigth lines correspond to model outcomes under various tax regimes, whereas the dotted line shows the efficient outcome. As is clear, the distribution of risks when $\tau = 0$ is too much skewed to the left in comparison to efficiency. As τ increases, the distribution becomes less tilted. At $\tau = 0.09$, the distribution is almost identical to the efficient distribution. When $\tau = \tau^z = 0.25$, we have the limit case, where all wages are equal and hence $f(\theta) \propto \frac{1}{h(\theta)}$.
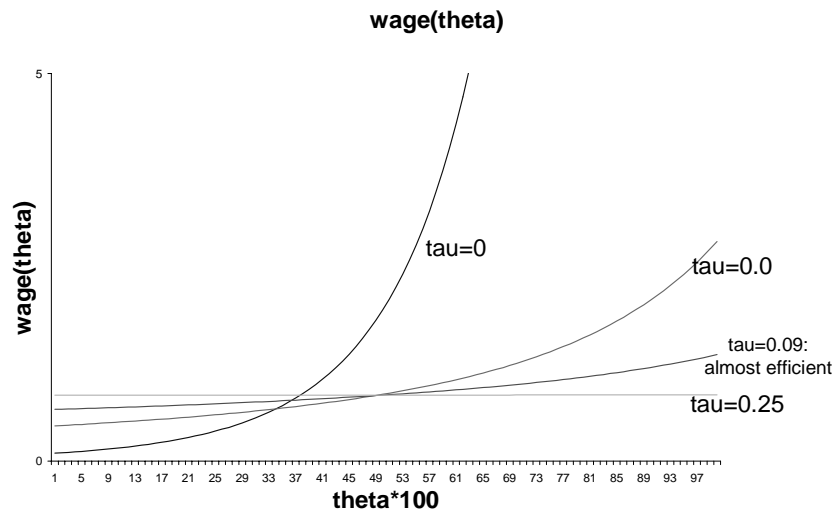
Figure 2: the relation between wage profiles and taxation

Figure 2 shows the wage profile of the same tax regimes. At $\tau = 0$, wages rise very fast with risks and hence with θ. As τ increase, the wage profile becomes less skewed to the right. At $\tau = 0.09$, we get the almost efficient wage curve. In the very limit case of $\tau = 0.25$, wages are constant.
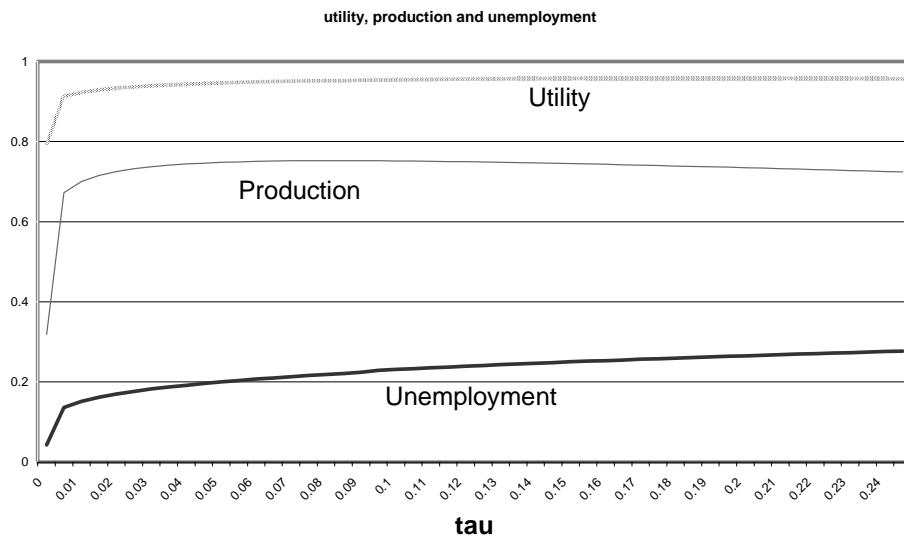

Figure 3: the effect of taxation on utility, production, and unemployment

Figure 3 shows the relation between τ one the one hand and utility, pro-

duction and unemployment. Each three rises quite rapidly for very low levels of τ. Production peaks very soon, i.e. at $\tau = 0.09$ and then slowly levels off. Utility peaks much later at $\tau = 0.22$, and unemployment increases till the limit of $\tau = 0.25$.

In this homogeneous-worker model there is a negative relationship between wages and risks: the higher the wages, the higher the risk of unemployment. This corresponds to the results of the classic hedonic wage literature that started with Rosen (1972) where individuals pay with lower wages for good amenities, in this case low risks of unemployment. This trade-off is also present in Acemoglu and Shimer (1999). It is empirically very implausible however that the better paid run more chance of being unemployed, except perhaps for some very high paying jobs such as CEO's. Daniel and Sofer (1998) show that in the majority of empirical studies, the usual finding is that the higher the wages, the better the amenities, counter to our simple model's prediction. Unemployment levels are for instance generally higher for individuals at the lower earning potential levels. Individual heterogeneity plays a major role here: those with lowest wages are generally also those with the lowest ex ante earning potential. There is hence a relation between individual abilities and wages that also impacts on the observed relation between risk and wages. We therefore augment the basic model with individual talent heterogeneity.

2.1   Including individual talent heterogeneity

Individuals have an innate talent $1 > q > 0$ that is drawn from a differentiable population distribution $G(q)$. This talent is interpreted as an efficiency unit. This means the wages of someone who works with talent q and choice of specialisation θ is $qw(\theta)$. We now conjecture that it will be the case that all individuals with a certain talent will choose a particular θ and that this implicit function $q(\theta)$ is either increasing or decreasing on its domain (a no-crossing conditions). This conjecture will have to be confirmed ex post. The efficiency units of labour supplied for speciality θ is then $g(q(\theta))\frac{dq(\theta)}{d\theta}q(\theta)h(\theta)$.[2] This means that total output can be written as

$$y = [\int_0^1 (\frac{dq(\theta)}{d\theta}g(q(\theta))q(\theta)h(\theta))^\gamma d\theta]^{1/\gamma}$$

[2]In the more general case, the number of efficiency units should be written as $\int_{x=q(\theta)} h(\theta)x(\theta)dG$.

11

with the wage per efficiency unit $w(\theta) = \frac{\partial y}{d[g(q(\theta))\frac{dq(\theta)}{d\theta}q(\theta)h(\theta)]} = (\frac{dq(\theta)}{d\theta}g(q(\theta))q(\theta)h(\theta))^{\gamma-1}*$ $y^{1-\gamma}$.

Individual utility maximization will now mean that at θ, an individual with quality $q(\theta)$ is indifferent in her choice:

$$\frac{\partial\{h(\theta)u((1-\tau)qw(\theta)) + (1-h(\theta))u(b)\}}{\partial\theta}|^{q(\theta)} = 0$$

For each quality level q, this requirement leads to an indifference curve on the $\{\theta, w\}$ space. Applying the argument in the seminal paper by Rosen (1972), this means that in equilibrium the wage curve $w(\theta)$ will be the envelope of these indifference curves, which in turn determines q(θ). In order to calculate the equilibrium in any practical instance, we note that we can simply trace the condition above for any w(0) to arrive at a w(θ), which in turn determines everything. We then pick that w(0) that is itself implied by the outcomes. Proposition 3 gives conditions for uniqueness and of the main relation of interest.

Proposition 3. *i) The model with individual heterogeneity has a unique $w(\theta)$, and hence a unique level of $E\{U(q)\}$. ii) The equilibrium is inefficient for any τ, though $\frac{dy}{d\tau}|^{\tau=0} > 0$ and $\frac{dU(q)}{d\tau}|^{\tau=0} > 0$. iii) Any continuous distribution of observed wages is supported. iv) Individuals with $q < \frac{b}{w(1)}$ are voluntarily unemployed. v) Total wages always increase with talent ($\frac{dwq}{dq} > 0$) but will only decrease with risk ($\frac{dwq}{d(1-h)} < 0$) when $\{u(x)-u(b)\}(-\frac{u''(x)}{u'(x)}x-1) + xu'(x) > 0$, which is the case whenever the degree of relative risk aversion is constant, or doesn't cross 1.*
Proof: see the Appendix.

We may note that $\frac{dwq}{d(1-h)} < 0$ holds for most popular utility functions, such as $u(w) = \ln(a+w)$ with a>0, $u(w) = w^{\alpha}$ with $0 < \alpha < 1$, and $u(w) = 1 - e^{-\alpha w}$ with $\alpha > 0$. The intuition is that individuals with high innate talents can 'buy' security by talking less risky professions that pay less per efficiency unit. Observed total wages ($= qw(q)$) are then increasing in talent but decreasing in risks. In such cases, the observed relations of risks and wages is spurious and due to innate talent heterogeneity. This is slightly different to the explanation given by Hwang et al. (1998) and Daniel and Sofer(1998). In their models, the same empirical prediction arose because workers with high bargaining power extracted surplus from efficient firms by having both

higher wages and better amenities. In Hwang et al. (1998) this bargaining power derived from the possibility of on-the-job-search in the presence of heterogeneity in the productivity of firms whereas in Daniel and Sofer (1998) the bargaining power derived from the presence of a union. In the model of this paper however, individual heterogeneity can cause a similar relationship but does not rely on the presence of a market-distortion such as some firms that are not on the production frontier or the presence of unions. Our explanation stresses unobserved differences in individual characteristics which increase both wages and the preference for amenities and is therefore more general. Because of the spurious relation between wages and risks on the individual level, we expect different relations between wages and risks on the aggregate than on the individual level. We can for instance refer to Van Vuuren and Van der Berg (2001) who find that risks and wages are positively related at the sector level in the Netherlands, whereas risks and wages are negatively related at the individual level in the Netherlands (e.g. Frijters et al. 2001).

The main prediction about education that comes out of these static models is that the proportion of individuals into less specialised occupations is higher in economies with lower levels of welfare. Now, the welfare system is more generous and elaborate in the European and richer Asian countries than it is in the US. Ashton and Greene (1996) extensively review the education systems in most OECD countries and in the fast growing Asian countries. They note that the education system in the US is indeed quite generalist. Secondary and tertiary education in the US is based on many subjects, and little specialisation takes place in formal education. Most of the European and Asian countries on the other hand, have education systems that are much more specialised. Tertiary education is much more focussed on a small subset of subjects, and even much of secondary education is occupation-specific. The extreme case is the vocational system in Germany were large proportions of the population learn only very specific skills.

3   Dynamics

We first extend the two-stage model above into an infinite horizon, continuous time OLG model, where we focus exclusively on steady states. In order to look at what happens after unanticipated shocks, we then set up an empirical version of the model which we calibrate.

3.1   An analytical dynamic model

Consider an infinite-period continuous overlapping generations model with a total measure of individuals equal to 1. Ex-ante homogenous individuals can be unemployed, employed and can die. At a mortality rate m individuals die, who are immediately replaced by a new individual. This new individual chooses her specialisation and then starts out being unemployed. There is a common discount rate ρ. The distribution of specialisation chosen by the individuals who enter into the economy at time s is denoted as $f_s(\theta)$. Calendar time is denoted by t. The job-arrival rate for the unemployed with specialisation θ is denoted as $\lambda(\theta) > 0$, and the job-destruction rate by $\delta(\theta) > 0$. We define specialisation as $\frac{d\lambda(\theta)}{d\theta} < 0$ and $\frac{d\delta(\theta)}{d\theta} > 0$.

The total measure of workers and unemployed with specialisation θ at time 0 will equal $\int_{-\infty}^{0} me^{ms} f_s(\theta) ds$, where me^{ms} is the density of the individuals alive at 0 who are born at time $s < 0$. The density of individuals with specialisation θ who entered at time s employed at time t>0 is then $p(\theta, t - s) f_s(\theta) me^{m(s-t)}$ where the probability $p(\theta, t - s)$ is defined by the differential equation

$$\frac{\partial p(\theta, t - s)}{\partial t} = \lambda(\theta)\{1 - p(\theta, e, t - s)\} - \delta(\theta) p(\theta, e, t - s)$$

and the initial condition $p(\theta, 0) = 0$. Solving this differential equation leads to

$$p(\theta, t - s) = \frac{\lambda(\theta)}{\lambda(\theta) + \delta(\theta)}(1 - e^{-(t-s)(\lambda(\theta) + \delta(\theta))})$$

This probability has standard properties: $\frac{\partial p(\theta, t-s)}{\partial t} > 0$, $\frac{\partial^2 p(\theta, t-s)}{\partial^2 t} < 0$, and $lim_{t \to \infty} p(\theta, t - s) = \frac{\lambda(\theta)}{\lambda(\theta) + \delta(\theta)}$. The total population measure of individuals with specialisation θ employed at time t is denoted as $G_t(\theta)$ and equal to $\int_{-\infty}^{t} p(\theta, t - s) f_s(\theta) me^{m(s-t)} ds$. Total production in each period is then

$$y_t = (\int \{G_t(\theta)\}^\gamma d\theta)^{\frac{1}{\gamma}}$$

and wages solve $w_t(\theta) = \frac{\partial y_t}{\partial G_t(\theta)}$.

We can now find the Euler equations for the value of unemployment and employment that characterise maximising behaviour. Denoting the expected utility value of unemployment as V^{UN} and the utility value of employment as V^{EM}, there holds

14

$$\begin{aligned}
(\rho + m)V^{UN} &= u(b) + \lambda(\theta)\{V^{EM} - V^{UN}\} \\
(\rho + m)V^{EM} &= u((1 - \tau)w(\theta)) + \delta(\theta)\{V^{UN} - V^{EM}\}
\end{aligned}$$

Solving these equations for V^{UN} yields

$$(\rho+m)V^{UN} = \frac{\rho + m + \delta(\theta)}{(\lambda(\theta) + \delta(\theta) + \rho + m)}u(b) + \frac{\lambda(\theta)}{(\lambda(\theta) + \delta(\theta) + \rho + m)}u((1-\tau)w(\theta))$$

or

$$(\rho + m)V^{UN} = (1 - a(\theta))u(b) + a(\theta)u((1 - \tau)w(\theta))$$

Looking only at RE stationary steady states, V^{UN} has to be the same for every θ. This means the solution has to satisfy

$$\frac{da(\theta)}{d\theta}[u((1 - \tau)w(\theta)) - u(b)] = -w'(\theta)(1 - \tau)a(\theta)u'((1 - \tau)w(\theta))$$

where $\frac{da(\theta)}{d\theta} < 0$. Writing the continuous time model in this way has hence given a solution equation that is virtually the same as in the static model, and for which the arguments in proposition 1 apply. The solution equation again maps out a unique $w(\theta)$ from any starting point $w(0)$. From $w(\theta)$, we can then map out a unique $\frac{G(\theta)}{G(0)}$, which in turn uniquely determines a static $f(\theta)$ and y. Because these in turn again lead to a unique $w(0)$, there is a closing equation that is continuous and which has to have at least one fixed point. Existence is thereby assured. The arguments on uniqueness are the same because it is again not possible to change $G(\theta)$ (via $f(\theta)$) without making some individuals strictly better and some others strictly worse off, which is not possible. Obviously, $w(\theta)$ increases in $\delta(\theta)$ and decreases in $\lambda(\theta)$.

When it concerns efficiency and the role of τ, this dynamic model is more problematic because there is no single risk parameter but there are now two, i.e. $\lambda(\theta)$ and $\delta(\theta)$. For efficiency to hold, it has to be the case that $\frac{\partial y}{\partial f(\theta)}$ is the same for all θ. Using the fact that $\lambda(\theta) = \infty$ for $\theta = 0$ (which is just a normalisation), there has to hold for an efficient $f(\theta)$ that $\frac{\partial y}{\partial f(\theta)} = f(0)^{\gamma-1}y^{1-\gamma} = w(0)$. Denoting the efficient solution as $f^e(\theta)$ and $w^e(\theta)$, there then holds $\frac{f^e(\theta)}{f^e(0)} = \bar{p}(\theta)^{\frac{\gamma}{1-\gamma}}$ where $\bar{p}(\theta) = \int_{-\infty}^{0} p(\theta, e, t-s)me^{ms}ds$.

Also, $\frac{dw^e(\theta)}{d(1-\bar{p}(\theta))} = \frac{w^e(\theta)}{\bar{p}(\theta)}$. We can hence examine the efficiency of the equilibrium by checking whether $w(\theta) = \frac{w(0)}{\bar{p}(\theta)}$ can optimize V^{UN}. First we calculate

$$
\begin{aligned}
\bar{p}(\theta) &= \int_{-\infty}^{0} \frac{\lambda(\theta)}{\lambda(\theta) + \delta(\theta)}(1 - e^{s(\lambda(\theta)+\delta(\theta))})me^{ms}ds \\
&= \frac{\lambda(\theta)}{\lambda(\theta) + \delta(\theta) + m}
\end{aligned}
$$

and

$$
\frac{d\bar{p}(\theta)}{d\theta} = \frac{\lambda'(\theta) - \lambda(\theta)\frac{\lambda'(\theta)+\delta'(\theta)}{\lambda(\theta)+\delta(\theta)+m}}{\lambda(\theta) + \delta(\theta) + m} < 0
$$

For efficiency to be possible, there would have to hold that

$$
\frac{dw(\theta)}{d(1-\bar{p}(\theta))} = \frac{-\frac{da(\theta)}{d\theta}[u((1-\tau)w(\theta)) - u(b)]}{a(\theta)(1-\tau)u'((1-\tau)w(\theta))} \Big/ \frac{d\bar{p}(\theta)}{d\theta} = \frac{(\lambda(\theta) + \delta(\theta) + m)w(\theta)}{\lambda(\theta)}
$$

which after some manipulations can be written as the condition that

$$
xu'(x) = A(\theta)\{u(x) - u(b)\}
$$

where

$$
\begin{aligned}
x &= (1-\tau)w(\theta) \\
A &= \frac{-\lambda'(\theta) + \lambda(\theta)\frac{\lambda'(\theta)+\delta'(\theta)}{\lambda(\theta)+\delta(\theta)+\rho+m}}{-\lambda'(\theta) + \lambda(\theta)\frac{\lambda'(\theta)+\delta'(\theta)}{\lambda(\theta)+\delta(\theta)+m}} > 0
\end{aligned}
$$

from which it is clear that, under any τ, this condition can only be satisfied by a linear utility function at particular parameter values. Hence, in general the equilibrium is again never efficient under any tax system.

In order to see whether we can follow the same arguments in the dynamic case as in the two-period case about the effect of increasing τ at $\tau = 0$, we can note that

$$
\frac{dw(\theta)}{d(1-\bar{p}(\theta))}\Big|^{\tau=0} > A(\theta)\frac{w(\theta)}{\bar{p}(\theta)}
$$

Now, in the case that $\rho\lambda'(\theta) < -\rho\delta'(\theta)$, then $A(\theta) > 1$ and hence $\frac{dw(\theta)}{d(1-\bar{p}(\theta))}\Big|^{\tau=0} > \frac{w(\theta)}{\bar{p}(\theta)}$ which corresponds to inefficiently low risk-taking and

16

whose characteristics would correspond to the 2-period case. Looking at this condition, this will hold whenever the negative change in job-finding rates is lower than the positive change in job-destruction rates. In these circumstances, the introduction of an UB would thus again improve efficiency. Given that in most cases, job-destruction rates are lower than job-creation rates (employment lasts longer than unemployment), the case that the changes in job-finding rates are more important, seems the generic case.

However, in the less likely case that $\rho\lambda'(\theta) > -\rho\delta'(\theta)$, there holds that $A(\theta) < 1$. For low degrees of risk-aversion, it then becomes possible that $\frac{dw(\theta)}{d(1-\bar{p}(\theta))}|^{\tau=0} < \frac{w(\theta)}{\bar{p}(\theta)}$ in which case individuals take too much risk, even at $\tau = 0$. Hence we cannot a priori say whether taxes will be output increasing or not. The reason behind this unexpected result is discounting: individuals attach more value to what happens in the near future than output maximisation dictates. Then, simply because job-destruction can only occur at a later date than getting a job, individuals attach a disproportionate importance to $\lambda'(\theta)$ and not enough to $\delta'(\theta)$. It is then possible that a too high proportion of individuals take specialisations with a high job-destruction rate from an output maximising point of view. Hence, surprisingly, whether aggregate choices are too risky or not depends on the precise shape of $\lambda(\theta)$ and $\delta(\theta)$, although in all the parametric examples looked at in this paper, $A(\theta) > 1$ (because in all applications $\delta'(\theta) = 0$ already).

3.2 A calibrated dynamic model

In order to find out what happens when shocks occur to the underlying parameters, we set up a calibrated version of the model to highlight the main dynamics at work.

We look at a discrete time model with two choices of specialisation, high and low. Ex ante identical individuals choose their specialisation at t=0, then enter unemployment. After 20 periods they die. The population grows at a rate n, meaning that the cohort that dies is replaced by a new cohort that is a fraction $(1+n)^{20}$ bigger. Because of the constant returns to scale production function, population growth will not affect anything in the steady state. Population growth will however serve to 'dampen out' fluctuations. Each period a fraction β_t of the new cohorts will choose general education, and a fraction $(1-\beta_t)$ chooses specialist education. Those with general education have job-finding rates equal to λ^H and job-destruction rates equal to δ^H. Those with specialist education have job-finding rates equal to $\lambda^S < \lambda^G$ and

job-destruction rates equal to $\delta^S \geq \delta^H$. Individuals are assumed to have rational expectations and have logarithmic utility functions: $u(x) = \ln(x + A)^\alpha$. Because we calibrate working lives, each period corresponds to about 2.5 years. Taking standard estimates for the discount rate from empirical studies (e.g. Frijters and Van der Klaauw 2001), we take ρ equal to 10% a year. There is no mortality before period 20.

We first show some baseline calculations of the steady state, where we construct two different baseline economies. The choice of the key variables $\{\gamma, \tau, n, \lambda^G, \lambda^S, \delta^G, \delta^S\}$ is calibrated on statistics from the US and the EU. In line with Sargent and Ljungqvist (1998), we do not allow job-destruction rates or basic production processes to differ. We have hence set $\{\gamma, \lambda^G, \lambda^S, \delta^G, \delta^S\}$ equal for both economies, in order to allow for proper comparisons of the dynamics in the shock experiments. Also, the parameters were set such that the level of specialists is much higher in in one economy, in order to be assured that our risk-arguments have some relevance for the two economy. The statistics we took into account are average job-finding rates, unemployment rates, population growth rates in the last 30 years, job-destruction rates, and level of UB. The underlying data is from the OECD.

Table with baseline functions.

Table 1: baseline callibrations for a high tax and a low tax economy.

	'EU'	'US'
Specific inputs		
n per year	0.01	0.02
τ	0.15	0.02
Outcomes		
β	0.59	0.98
$w^G * (1 - \tau)$	0.89	0.98
$w^S * (1 - \tau)$	0.96	1.50
benefits	0.80	0.55
Average production	0.89	0.97
unemployment	0.17	0.04
utility	-0.07	0.29

Common inputs: A=0.1, ρ per year=0.1, $\gamma = 0.9$, $\lambda^G = 0.9$, $\lambda^S = 0.3$, $\delta = 0.2$

In the 'EU' baseline model, we see that taxes and unemployment levels are higher than in the 'US' baseline model. Welfare is slightly higher in the 'US'. In the 'US', output (per hour input), population growth, and wage-differentials are also higher. The sensitivity to taxes is quite large: even

though basic production possibilities are set equal for both economies, the differences in the taxes translate to big differences in unemployment, though the differences in benefits are not that large.

We now turn to the effect of unanticipated shocks. Starting from steady state, we perturb the economy at t=0. We then assume that expectations instantaneously realign themselves, i.e. after the shock all individuals know what will happen next. We look at three different shocks:
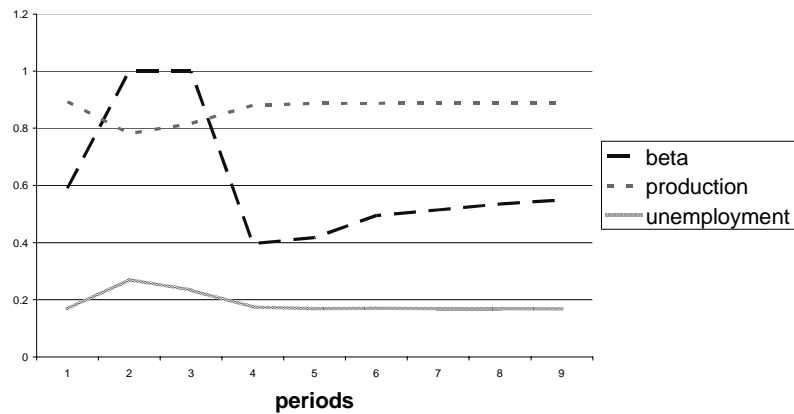
1. A recession. At t=0, δ is equal to 0.5, and λ^G and λ^S are halved. At t=1 all parameters return to baseline.

2. A permanent biased search shock. At t=0, $(1-\lambda^S)$ is halved.

3. A welfare system change. At t=0, τ is halved.

We show the results in graphs.

Response of the 'US' to an unemployment shock
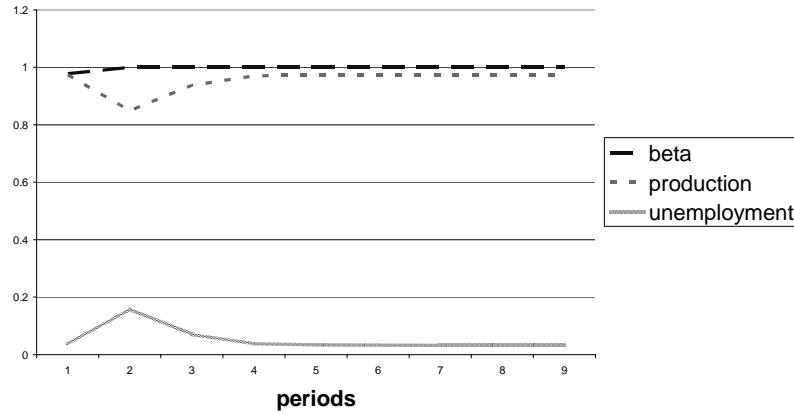

Response of the 'EU' to an unemployment shock

The effect of a recession corresponds somewhat to the historical data on unemployment dynamics after the oil shocks of the 70's (Blanchard en Wolfers, 1999): there is a similar initial drop in employment and output, but in the 'EU' the recovery in production and output takes about twice as long because the number of specialists is much higher. The different experience are especially pronounced when we look at the development of β : the recession temporarily increased the value of being a generalist. In the 'US', where most were generalists to begin with, this changed little, but in the 'EU', this led to big increases in β, which themselves created a 'ripple-effect' in subsequent periods. For the aggregate though, this means that the levels of return to work are smaller in the 'EU'. This is another explanation for the sluggishness

20

with which the EU unemployment rates came down after the oil price shocks in the 70's (Blanchard en Wolfers, 1999). Sargent and Ljungqvist (1998) explained this same phenomenon by arguing that unemployment benefits increased the willingness of individuals to wait for better jobs. Sargent and Ljungqvist combined this with the assumption that individual loose skills in unemployment. Hence unemployment benefits aggravated the negative shocks in their model because the high unemployment benefits indirectly adversely changed the characteristics of individuals whilst they were already on the labour market. The difference with this paper is that in this paper, there is no resort to a reduction in skills during unemployment, but an effect of unemployment benefits on the long term composition of the workforce.

In a sense, this calibrated model is set up such that the recovery is relatively quick because new cohorts can immediately enter the labour force. If we would allow for a period between the choise of education and entering the labour market, the recovery of areas with many specialists would take much longer.

**Response of the 'US' to decreased risks for specialists**
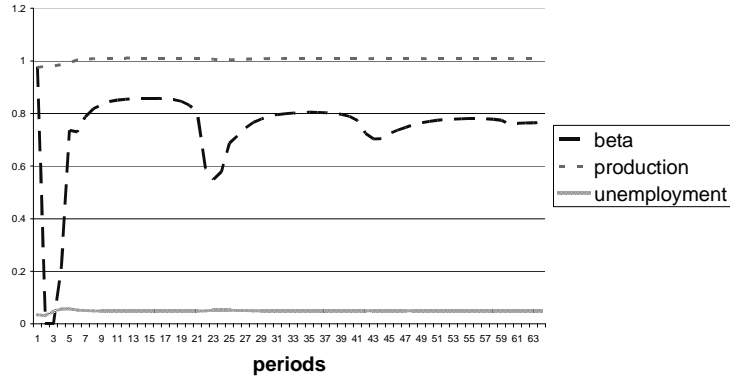
**Response of the 'EU' to decreased risks for specialists**
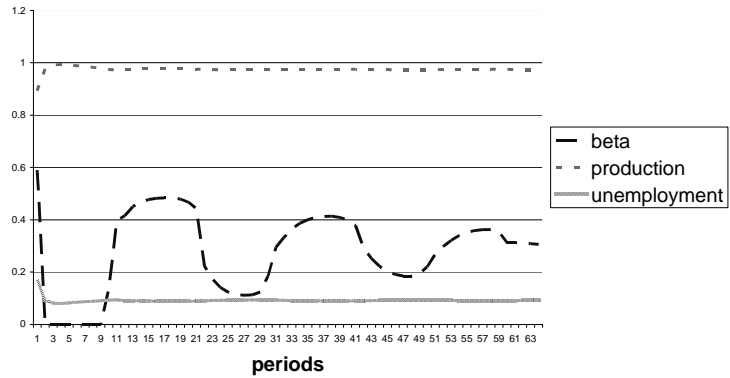
A permanent biased search shock has stronger effects on the 'EU' than on the 'US', simply because the proportion of specialists at the time of the shock is much higher in the 'EU'. The risks associated with being a specialist go down, which encourages more specialists. As an immediate effect, entire cohorts become specialists in the 'US' and the 'EU'. This causes a ripple effect for the subsequent periods that slowly dies down. In the long run, the proportion of specialists goes up in both economies. Also, production increases. Unemployment barely changes in the 'US'. In the 'EU', there is an initial sharp drop in unemployment as more specialists find a job. In the longer run however, the increased proportion of specialists, who still run higher risks than the generalists, means that the initial employment gains are reduced.

Response of the 'US' to a halving of tax rates
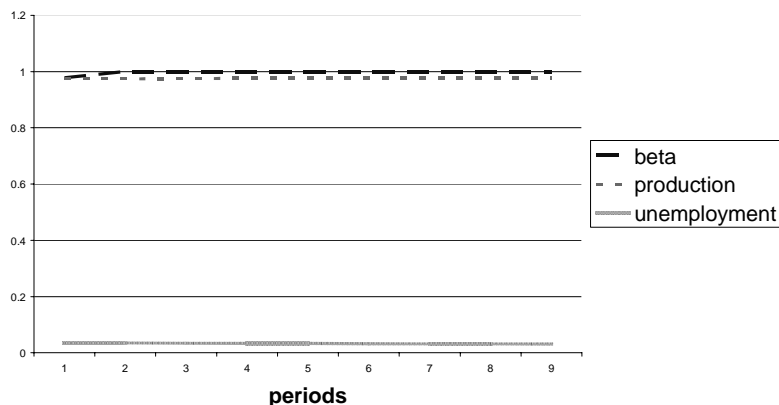
Response of the 'EU' to a halving of tax rates
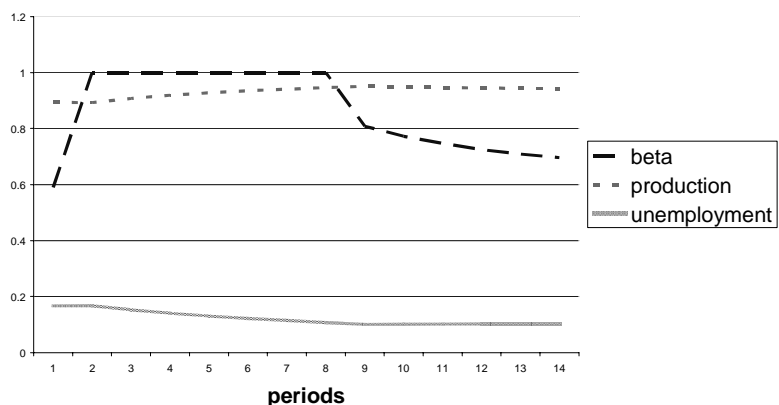
The welfare system change increases the incentives to become generalist. Because in the 'US', the proportion of generalists was already very high, the effect of lowering τ on unemployment and β is small: production and unemployment decrease very slightly. All that really changes is a large increase in the wages of specialists (not shown). There are big effects in the 'EU', where the changes in β are quite large, leading to long-lasting ripples. The effect on production and output is very delayed though, simply because it takes time for new cohorts of generalists to come through the education system and substantially change the composition of the labou force: it takes about 8 periods ($\approx$20 years) before production and unemployment have reached their new steady state. The long-term effects are that production increases and unemployment decreases because of the reduced numbers of specialists in the

23

economy. These predictions seem to mirror the sluggishness with which unemployment levels have been found to react to changes in the level of UB, and indeed to other welfare changes (Blanchard en Wolfers 1999, 2001; Dolado et al. 1996; Gruber and Wise 1997).

4   Conclusion and discussion

In this paper it is argued that unemployment benefits increase the incentives of individuals to make riskier educational investments, which at low unemployment benefit levels increases expected welfare, total output and unemployment. An area with high unemployment benefits therefore also has a higher number of specialists who have lower job-finding rates. This in turns makes such an area more vulnerable to general shocks, because unemployment levels return less quickly back to their natural rate. Also, increasing unemployment benefits increases unemployment, but with a large delay because it takes time before the more specialist newer cohort who are more frequently unemployed appear on the labour market. This is one explanation for the lack of responsiveness in unemployment rates that is frequently found for changes in unemployment benefit levels (e.g. Dolado et al, 1996; Blanchard and Wolfers 2001). It also explains why the US unemployment rates returned faster to a lower level after the 1970's oil price shocks, than did their European counterparts.

Apart from UB, there are several other risk-pooling systems. Disability benefits, early retirement benefits, and social welfare are obvious examples. If there are early life choices that affect the risks of these occurrences, then the existence of this risk-pooling may lead to exactly the same dynamics as UB does for unemployment: they will create delays between policy changes and changes in average outcomes, and they may affect both the distribution and efficiency of actual outcomes.

When there are several risk-pooling systems for the same risks, the introduction or expansion of one system may well affect investments into other systems. Maintaining close family and community ties for instance has as a likely benefit that one can count on support in the event of unemployment or other financial setbacks. In this sense, the maintenance and development of certain forms of social ties is a form of risk sharing. Welfare benefits then change the incentives for investments in these social ties and hence change the 'social fabric' of an economy.

Finally, the importance of long-term composition effects and general equilibrium effects found in this paper casts doubt upon the usefulness of looking at the partial effects of changes in policy on individuals with given early life choices.

Literature

1. Acemoglu, D., and R. Shimer (1999), 'Efficient unemployment insurance', *Journal of Political Economy* 107, pp. 893-928.

2. Acemoglu, D., and R. Shimer (2000), 'Productivity gains from unemployment insurance', *European Economic Review* 44, pp. 1195-1224.

3. Ashton, D. N., Green, F. (1996), *Education, training and the global economy*. Cheltenham, U.K.: Elgar

4. Aumann, R.J., (1977), 'The St. Petersburg Paradox: A Discussion of Some Recent Comments', *Journal of Economic Theory,* 14(2), pp. 443-45.

5. Blanchard, O.J., Wolfers, J. (1999), 'The Role of Shocks and Institutions in the Rise of European Unemployment: The Aggregate Evidence', *NBER working paper* 7282.

6. Blanchard, O.J., Wolfers, J. (2001), 'The Role of Shocks and Institutions in the Rise of European Unemployment: The Aggregate Evidence', *Economic Journal,* 110, pp. C1-33.

7. Bonnal, L., Fougere, D., Serandon, A. (1997), 'Evaluating the impact of french employment policies on individual labour market histories', *Review of Economic Studies*, 64, pp. 683-713.

8. Burdett, K. (1979), 'Unemployment insurance payment as a search subsidy: a theoretical analysis', *Economic Inquiry*, 42, pp. 333-343.

9. Daniel, C., Sofer, C. (1998), 'Bargaining, compensating wage differentials, and dualism of the labor market: theory and evidence in France', *Journal of Labor Economics*, 16, pp. 546-75.

10. Dolado, J., Kramarz, F., Machin, S., Manning, A., Margolis, D., Teulings, C., (1996), 'The economic impact of minimum wages in Europe', *Economic Policy*, 23, pp. 317-372.

11. Fredriksson, P, Holmlund, B. (2001), 'Optimal unemployment insurance in search equilibrium', *Journal of Labor Economics* 19, pp. 370-99.

12. Gruber, J. (1997), 'The consumption smoothing benefits of unemployment insurance', *Amerian Economic Review* 87, pp. 192-205.

13. Hopenhayn, G. and J.P. Nicolini, (1997), 'Optimal unemployment insurance', *Journal of Political Economy* 105, pp. 412-438.

14. Hwang, H., Mortensen, D.T., Reed, W.R. (1998), 'Hedonic wages and labor market search', *Journal of Labor Economics* 16, pp. 815-847.

15. Hwang, H., Reed, W.R., Carlton, C. (1992), 'Compensating wage differentials and unobserved productivity', *Journal of Political Economy* 100, pp. 835-58

16. Ljungqvist, L., Sargent, T.J. (1998), 'The European unemployment dilemma', *Journal of Political Economy*, vol. 106(3), pp. 514-50.

17. Marimon, R., Zilibotti, F. (1999), 'Unemployment vs. mismatch of talents: reconsidering unemployment benefits', *Economic Journal* 109, pp. 266-291

18. Mortensen, D., Pissarides, C. (1999), 'Unemployment response to skill-biased shocks: the role of labour market policy', *Economic Journal* 109, pp. 242-265.

19. Petrongolo, B. (2001), 'Reemployment probabilities and returns to matching', *Journal of Labor Economics* 19, pp. 716-41.

20. Rosen, S. (1972), 'Hedonic prices and implicit markets: product differentiation in pure competition', *Journal of Political Economy* 82, pp. 34-55.

21. Van den Berg, G.J., and A.P. van Vuuren (2001), 'The effect of search frictions on wages', Working Paper, Free University, Amsterdam.

Appendix

First, a useful result:

Lemma 1. *When going from an initial distribution $F_1(\theta)$ to a new distribution $F_2(\theta)$ with a continuous CES, wages will be higher for at least one θ and wages will be lower for at least one θ.*

Proof: consider the level $\theta^* = \arg_\theta \max\{\frac{f_2(\theta)}{f_1(\theta)}\}$. If $F_2 \neq F_1$, then there has to hold that $\frac{f_2(\theta^*)}{f_1(\theta^*)} > 1$. There holds

$$
\begin{aligned}
\frac{w_2(\theta^*)}{w_1(\theta^*)} &= \frac{(h(\theta^*)f_2(\theta^*))^{\gamma-1}}{(h(\theta^*)f_1(\theta^*))^{\gamma-1}} \frac{\left[\int (h(\theta)f_2(\theta))^\gamma d\theta\right]^{\frac{1}{\gamma}-1}}{\left[\int (h(\theta)f_1(\theta))^\gamma d\theta\right]^{\frac{1}{\gamma}-1}} \\
&= \frac{(f_2(\theta^*))^{\gamma-1}}{f_1(\theta^*)} \cdot \left[\frac{\int (h(\theta)f_2(\theta))^\gamma d\theta}{\int (h(\theta)f_1(\theta))^\gamma d\theta}\right]^{\frac{1}{\gamma}-1}
\end{aligned}
$$

Now,

$$
\begin{aligned}
\int (h(\theta)f_1(\theta))^\gamma d\theta &= (\frac{f_2(\theta^*)}{f_1(\theta^*)})^{-\gamma} * \int (h(\theta)\frac{f_2(\theta^*)}{f_1(\theta^*)}f_1(\theta))^\gamma d\theta \\
&> (\frac{f_2(\theta^*)}{f_1(\theta^*)})^{-\gamma} \int (h(\theta)f_2(\theta))^\gamma d\theta
\end{aligned}
$$

Hence

$$
\frac{w_2(\theta^*)}{w_1(\theta^*)} < \frac{(f_2(\theta^*))^{\gamma-1}}{f_1(\theta^*)} \left[(\frac{f_2(\theta^*)}{f_1(\theta^*)})^\gamma\right]^{\frac{1}{\gamma}-1} < 1
$$

Which implies that $w_2(\theta^*) < w_1(\theta^*)$. Using the same argument for $\theta^* = \arg \min\{\frac{f_2(\theta)}{f_1(\theta)}\}$, we can see that there is also a θ^* for which $w_2(\theta^*) > w_1(\theta^*)$. When $f(\theta)$ is continuous, the strict inequality implies that any change in $f(\theta)$ means that a whole range of wages must decrease and another range of wages must increase.

Proof of proposition 1.

Uniqueness of w(θ). First, we note that the differential equation

$$
-h'(\theta)(u(w(\theta)) - u(0)) = (1 - \tau)w'(\theta)h(\theta)u'(w(\theta))
$$

defining the equilibrium $w(\theta)$ is well-behaved in the sense that $w(\theta)$ is uniquely determined by a $w(0)$. Also, $\frac{\partial w(\theta)}{\partial w(0)}$ is continuous and bigger than 0. In turn, the equation $\frac{w(\theta)}{w(0)} = \frac{(h(\theta)f(\theta))^{\gamma-1}}{(h(0)f(0))^{\gamma-1}}$ means that $\frac{w(\theta)}{w(0)}$ uniquely determines $\frac{f(\theta)}{f(0)}$ and $f(0)$ is then solved by $\int f d\theta = 1$. The function $f(\theta, w(\theta))$ is also continuous in $w(\theta)$. By implication of $w(\theta)$ being continuous in $w(0)$, the implicit function $f(\theta, w(0))$ must therefore also be continuous in $w(0)$.

Finally, equilibrium requires that the level $w(0)$ also solves

$$(h(0)f(0, w(0)))^{\gamma-1} \left[\int (h(\theta)f(\theta, w(0)))^{\gamma} d\theta \right]^{\frac{1}{\gamma}-1} = w(0)$$

Now, for $w(0) \downarrow 0$, we first can note that at $\lim_{w(0)\downarrow 0}$ the defining condition $-h'(\theta)(u(w(\theta)) - u(0)) = (1 - \tau)w'(\theta)h(\theta)u'(w(\theta))$ reduces to $\frac{w'(\theta)}{w(\theta)} = \frac{-h'(\theta)}{h(\theta)}$. Translating this into $f(\theta)$, implies that the left-hand side of the above expression will converge to some positive number. For $w(0) \uparrow \infty$, we can note that that in the limit, $u'(w(\theta))$ becomes constant, which in turn again pins down the the left-hand side of the expression above to a finite number. Because of the continuity of $f(\theta, w(0))$ the fixed-point theorem hence applies and there must be at least some level $w(0)$ for which the condition is satisfied.

Considering uniqueness, suppose there are two wage function w_1 and w_2 that are an equilibrium. Without loss of generality, take $w_2(0) > w_1(0)$. Then, it has to be the case for w_2 that all will prefer $\theta = 0$ above any other level θ unless $w_2(\theta) > w_1(\theta)$ for any $\theta > 0$. If this does not hold, there is a positive mass of quality that will choose 0 and there will be no mass choosing a quality slightly above 0. This in turn would mean $w_2(0) = 0$, which can't be the case. Hence, there can only be a second equilibrium if $w_2(\theta) > w_1(\theta)$ for all $\theta > 0$. As we know from lemma 1, this is an impossibility. If there is an equilibrium, it hence has to be unique.

In this proof it was not necessary to assume that $w'(\theta)$ or $f(\theta)$ is continuous. This means we can include cases where $h(\theta)$ is not-continuous. In such cases, we can apply the same proof, but simply note that $w(\theta)$ is then defined by the more general requirement that $E\{U(\theta)\} - E\{U(0)\} = 0$. In the text all the formulas are given for the continuous case though for ease of exposition.

Proof of proposition 2.

Our strategy for proving i) is to first prove that when taxation increases from 0 to some arbitrarily small ε, that there will be a $1 > \theta^* > 0$ for which $w(\theta)$ remains constant. Then, we will show that for all $\theta < \theta^*$, $w(\theta)$ has increased. For all other θ, $w(\theta)$ has decreased. Because $w'(h)$ will still be greater than $\frac{w}{h}$, we can then write the changes in $w(\theta)$ as a succession of production-increasing changes, which establishes production improvements. For θ^* we then shown that his utility has increased, which implies it has to have increased for all choices.

Looking at $\frac{d\frac{dw}{d(1-h)}}{d\tau}$ there holds

$$\frac{d\frac{dw}{d(1-h)}}{d\tau} = \frac{1}{(1-\tau)hu'((1-\tau)w(h))} * [\{-w(h)u'((1-\tau)w(h)) - \frac{db}{d\tau}u'(b)\} +$$
$$\{u((1-\tau)w(h)) - u(b)\} * \{\frac{1}{1-\tau} + \frac{w(h)u''((1-\tau)w(h)))}{u'((1-\tau)w(h))}\}]$$

Using that $\frac{db}{d\tau}|^{\tau=0} = \frac{\int hfwd\theta}{\int(1-h)fd\theta} > \frac{\int h(0)w(0)}{\int(1-h)fd\theta} > \frac{h(0)w(0)}{1-h(1)} > w(0)$ and that $u((1-\tau)w(\theta)) - u(b) < ((1-\tau)w(\theta) - b)u'(b)$ we hence know that $\frac{d\frac{dw}{d(1-h)}}{d\tau} < 0$ at $\theta = 0$ for τ close to 0. This in turn establishes that the whole wage function must have changed. We denote the changed wage function as $w(\theta, \varepsilon)$. Using the same argument as in Lemma 1, we thus know that there has to be a whole range of θ for which $w(\theta)$ has increased. Because of continuity, this also means there will be a $1 > \theta^* > 0$ for which $w(\theta, \varepsilon) = w(\theta)$. Because $w(\theta, \varepsilon)$ equalizes utility, this also implies that $sign\{w(\theta, \varepsilon) - w(\theta)\} = sign\{\theta^* - \theta\}$: because the importance of the risk of unemployment has decreased, taking more risks must have become relatively more attractive. This in turn means that $\frac{d(f(\theta,\varepsilon) - f(\theta))}{d\theta} > 0$. For small ε we can write $y(\tau = \varepsilon) - y(\tau = 0) = \int \{f(\theta, \varepsilon) - f(\theta)\}h(\theta)w(\theta)d\theta$. Now, because $\frac{dw}{d(1-h)} = \frac{u((1-\tau)w(h)) - u(b)}{(1-\tau)hu'((1-\tau)w(h))}$ is continuous in τ, $\frac{dw}{d(1-h)}$ will still be greater than $\frac{w}{h}$ for very small τ. Hence we can still use that $\frac{dh(\theta)w(\theta)}{d\theta} > 0$, implying that $y(\tau = \varepsilon) - y(\tau = 0) > 0$.

Now, using that $\frac{dw}{d(1-h)}|^{h=1} / \frac{d\frac{dw}{d(1-h)}}{d\tau}|^{\tau=0,h=1} < -1$ and that average wages have increased, we also know that $h(\theta^*) < (1 - \varepsilon)$ and thereby that $w(0, \varepsilon) > (1 + \varepsilon)w(0)$. This in turn means that the wage increase at $\theta = 0$ more than offsets the tax increase, which establishes a utility increase for all θ.

ii) First, we simply define τ^z as the τ that solves $(1-\tau)w(0, \tau) = b(\tau)$. For this level of τ, $\frac{dw}{d(1-h)}$ has to equal 0 because otherwise someone could increase their expected utility by changing θ. Hence $w(\theta) = w(0)$ and $f(\theta) \propto h(\theta)^{-1}$.

Also, $\frac{d\frac{dw}{d(1-h)}}{d\tau}\big|^{\tau=\tau^z} > 0$. This in turn would imply that if $\tau > \tau^z$, that $\frac{dw}{d(1-h)} < 0$ and that individuals would want to take as much risk of unemployment as possible, which in turn means they would not want to work at all.

Proof of proposition 3.

i) Existence. We can follow the proof of the homogenous case by noting that the main solution equation uniquely tracks out $w(\theta)$ as a function of $w(0)$. What is now uniquely defined by $\frac{w(\theta)}{w(0)}$ is $\frac{\frac{dq(\theta)}{d\theta} g(q(\theta))q(\theta)h(\theta)}{\frac{dq(0)}{d\theta} g(q(0))q(0)h(0)}$ and it is now the function $\int^R x(\theta)dG(x = q(\theta))$ that is uniquely, continuously and implicitly defined by $w(0)$ and by the requirement that $\{q(\theta)\} = \{0 \le x \le 1\} \cup \{qw(\theta) \ge b\}$. Voluntary nuemployment arises for those q where $qw(1) < b$.

Uniqueness. Suppose the equilibrium is not unique. Without loss of generality, take $w_2(0) > w_1(0)$. Then, it has to be the case that all with quality q in a small region near $q(0)$ will prefer $\theta = 0$ above any other level θ unless $w_2(\theta) > w_1(\theta)$ for $\theta > 0$ also. If this does not hold, there is a positive mass of quality that will choose 0 and there will be no mass choosing a quality slightly above 0. This would in turn invalidate the initial assumption. By forward induction, there can hence only be a second equilibrium if $w_2(\theta) > w_1(\theta)$ for all $\theta > 0$. As we know from lemma 1, this is an impossibility. If there is an equilibrium, it therefore has to be unique.

iii) and iv) The arguments on taxation and the observed wage distribution trivially carry over from the homogeneous case.

v) First we will prove that it can never be the case that $\frac{dwq}{dq} < 0$. For this we note that $E\{U(q_1)\} > E\{U(q_2)\}$ when $q_1 > q_2$. It thus has to be the case that $h(q_1)\{u((1-\tau)q_1 w(q_1)) - u(b)\} + u(b) > h(q_2)\{u((1-\tau)q_2 w(q_2)) - u(b)\} + u(b)$. Now, because of utility maximisation, we also know that $h(q_1)\{u((1-\tau)q_2 w(q_1)) - u(b)\} + u(b) \le h(q_2)\{u((1-\tau)q_2 w(q_2)) - u(b)\} + u(b)$. Substracting the second inequality from the first, we get $h(q_1)\{u((1-\tau)q_1 w(q_1)) - u((1-\tau)q_2 w(q_2))\} > 0$. In turn, this means that $q_1 w(q_1) > q_2 w(q_2)$. Hence we indeed know that $\frac{dwq}{dq} > 0$, i.e. total wages will be higher for individuals with higher talents. We then know that $\frac{dwq}{d(1-h)} < 0$ iff $\frac{d^2h}{dwdq} > 0$ because then individuals with higher talent have a greater preference for less risk than those with less

talent. There now holds that

$$\frac{d^2h}{dwdq} = u'(x)\frac{\{u(x) - u(b)\}(-\frac{u''(x)}{u'(x)}x - 1) + xu'(x)}{\{u(x) - u(b)\}^2}$$

with $x = (1 - \tau)w(h)$. Noting that $\frac{-u''(x)}{u'(x)}x$ is the degree of relative risk aversion ($=\sigma$), it immediately holds that $\frac{d^2h}{dwdq} > 0$ if $\sigma \geq 1$. Also, because $\{u(x) - u(b)\} \leq xu'(x)(1 - x\min_{b<a<x}\{\frac{u''(a)}{u'(x)}\})$, the condition will also hold when $\frac{-u''(x)x}{u'(x)}$ is constant or always less than 1.