

# Social Interaction in Regional Labour Markets <sup>\*</sup>

Jörg Heining<sup>†</sup> and Jörg Lings<sup>‡</sup>

15th April 2005

## Abstract

Social interaction, i.e. the interdependence of agents' behaviour via non-market activities, has recently become a focus of economic analyses. Especially in labour economics, social interaction has been used to explain various labour market outcomes. An important result of this literature is the proposition that labour markets are characterised by multiple equilibria. Thus, social interaction is used as an explanation for regional unemployment disparities. Building on this, we construct a Pissarides (2000) type search model with social interaction. Despite social interaction, this type of model is characterised by only one stable equilibrium. Using a unique data set on un-/employment spell data for Germany we analyse whether there exists multiple equilibria in regional labour markets. After controlling for structural differences we are able to show that the data supports the assumption of a unique equilibrium. As such, social interaction cannot explain regional unemployment disparities.

JEL: J6, R1.

Keywords: Social Interaction, Search Frictions, Survival Analysis, Regional Unemployment

---

<sup>\*</sup>For helpful comments and discussion we would like to thank, without implicating, Alfred Hamerle, Jürgen Jerger, Christoph Knoppik, Jochen Michaelis, Walter Oberhofer, Andrea Schrage and participants at the lunch seminar at the University of Regensburg.

<sup>†</sup>University of Regensburg, Joerg.Heining@wiwi.uni-regensburg.de, Tel.: ++49 941 943 2720

<sup>‡</sup>University of Regensburg, Joerg.Lings@wiwi.uni-regensburg.de, Tel.: ++49 941 943 2722

# 1 Introduction

Differences in the labour market performance (i.e. unemployment) in a cross-section of countries are usually being explained by the different institutional settings between these countries, see e.g. Layard and Nickell (1999). However, we do not only witness differences in unemployment rates between countries, but also within countries, i.e. at the regional level. These disparities in regional unemployment rates are large and persistent (see, e.g. Elhorst (2003)) and we cannot explain them by institutional differences since labour market institutions such as unemployment benefit legislations usually do not differ within a jurisdiction.

One possible reason for these regional differences in unemployment rates which has come under closer scrutiny is the notion of social interaction between agents at the regional level. Social interaction means that there exists some (non-market) interdependence between agents (e.g. preference based or constraint based interaction, see Manski (2000)) which shapes agents' behaviour and thus, influences the labour market performance.

The classical reference of the labour market effects of preference based social interaction is Diamond (1982). In this paper the agent's utility of engaging in production is a function of the other agents' engagement since this increases the probability of finding a suitable trading partner. More recent references which analyse the employment effects of preference based interaction are Lalive and Stutzer (2004) and Hedström et al. (2003).

Lalive and Stutzer (2004) argue that there exists an attitude towards the acceptability of living on the dole. This social norm exerts pressure on the unemployed and thus influences the unemployment rate. The strength of this norm, however, might differ across regions which leads to regional unemployment differentials. They present convincing evidence for Switzerland that there exists a correlation between the strength of this social norm and regional labour market performance.

Hedström et al. (2003) put forward very similar arguments. Their notion of social interaction is that the costs of being unemployed depends on the unemployment rate of the region one lives in. This is e.g. due to the fact that e.g. stigmatisation in a high unemployment environment is lower. As such, the utility of being unemployed is high and the pressure for finding a new job is low.

In contrast to these preference based interaction approaches, Topa (2001) and Conley and Topa (2002) put emphasize on constraint based social interaction. When searching for new jobs unemployed agents face the constraint of finding a suitable job. This constraint, however, is affected by the social environment the agent lives in since the majority of jobs is allocated via in-

formal channels. Thus an agent will c.p. find it easier to get a new job if she lives in a low unemployment environment and vice versa. Similar reasoning is applied in Selod and Zenou (2003) in which the probability of finding a job is a function of the social network one lives in.

An important feature common to many of the models which incorporate social interaction is the existence of multiple equilibria. Regions which are identical in their economic structure (productivity, educational structure and so on) thus could experience different labour market outcomes.<sup>1</sup> E.g. in the Hedström et al. (2003) framework (structurally) identical regions could implicitly coordinate on different labour market equilibria.

The question whether the root of regional unemployment disparities is the multiplicity of equilibria or structural economic differences is highly important for policy making. Multiple equilibria offer scope for governmental intervention to coordinate regions on the pareto dominating labour market equilibrium. In contrast the "structuralist" view of regional labour market disparities would not be that much in favour for regional policies.

Thus, since incorporating social interaction does not only increase our understanding of the functioning of labour markets, but also might be of policy relevance the question begs if social interaction inevitable results in multiple labour market equilibria and if the data support the view of multiplicity of labour market equilibria. This is the point of departure of the paper at hands.

We amend a Pissarides (2000) type search model of the labour market with social interaction. We model social interaction as a leisure externality. The idea is that unemployed agents have to invest time into the search process. The opportunity costs of this time investment depend on the time investment of the other agents in the region. This is due to the fact that agents like to spend their leisure time together. Thus, the utility from consuming leisure time will be a function of the leisure time of the other individuals. We demonstrate that although this type of social interaction might lead to self enforcing processes, the labour market equilibrium in the economy is unique and only driven by structural parameters of the model. Thus, social interaction cannot explain regional differences in unemployment rates.

In a next step we analyse empirically the possibility of multiple regional labour market equilibria using a unique micro-level data set on unemployment spell data for Germany. We estimate the hazard rate for leaving unemployment, i.e. the probability of leaving unemployment in the next instant

---

<sup>1</sup>Glaeser and Scheinkman (2000) analyse conditions under which preference based social interaction generates multiple equilibria. Their model, however, is very general and thus, does not directly address problems of search unemployment.

of time conditional on being unemployed, from 2000-2001. Controlling for structural individual and regional heterogeneity we find that the hazard rates do not significantly differ between regions. This supports our view that regional labour markets are characterised by a unique equilibrium and that regional differences in the labour market performance are due to economic differences between individuals living in structurally different regions.

The rest of the paper is organised as follows. The next section derives the first order conditions for the behaviour of individual agents and firms. These building blocks are put together in section 2.5 to analyse a symmetric general equilibrium. The empirical analysis follows in section 3. Firstly, we briefly describe the data set and the empirical model. Secondly, we present regression results and analyse the regional distribution of the hazard rates. The last section, eventually, summarises our results and concludes.

## 2 The Theoretical Model

### 2.1 Preliminaries

The basic framework of our analysis is the Pissarides (2000) model of frictional unemployment. In this model, (unemployed) workers are searching for jobs and firms are trying to fill vacancies. Both sides of the market are matched via a matching function. The rate  $m$  at which matching takes place depends positively on aggregate (=average) search intensity  $s$ , and the aggregate rates of unemployment  $u$  and vacancies  $v$ . We employ the familiar Cobb-Douglas specification of the matching function:

$$m = ((u)^{1-\alpha}(v)^\alpha)s. \quad (1)$$

Search intensity could be interpreted as input-augmenting efficiency of the matching function.<sup>2</sup> The probability of an efficiency unit of unemployment to be matched to a vacant job is given by:

$$\frac{m}{su} = \theta^\alpha, \quad (2)$$

where  $\theta$  (defined as  $v/u$ ) reflects labour market tightness.

---

<sup>2</sup>A quite similar idea is found, e.g. in Hosios (1990), where it is assumed that search intensity increases the probability of being matched. This will be exactly the case when the number of job matches is increasing in search intensity. We could also assume search intensity to be "unemployment" augmented. This would not change the main results.

## 2.2 Individuals' Behaviour

An unemployed individual decides on how much search intensity to "invest". This decision has to be based on the subjective values of being employed or unemployed. The Bellman equations for these states are given by:

$$rW_i = w_i + \lambda(U_i - W_i) \quad (3)$$

$$rU_i = \max_{s_i} \{b + l_i^\sigma + q_i(W_i - U_i)\} \quad (4)$$

where  $W_i$  is the value of being employed,  $U_i$  is the value of being unemployed,  $r$  is the rate of time preference (which is equal to the interest rate in a steady state),  $w_i$  is the wage the individual earns if employed,  $\lambda$  is the (exogenous) probability of destruction of a matched job and  $q_i$  is the individual probability of finding a job.

We assume risk neutrality on the part of the worker, i.e. utility is linear in the unemployment benefit  $b$ . The second argument in the flow utility function of unemployed workers captures the novel aspect in this framework:  $l_i$  is the amount of leisure the unemployed individual consumes. Individual leisure exhibits decreasing marginal utility (the utility function thus is quasi-linear). Since  $l_i = T - s_i$  (so searching a job is only costly in terms of time, i.e. in terms of foregone utility), the marginal costs of searching are increasing.<sup>3</sup>

In addition to individual leisure  $l_i$ , also aggregate leisure  $l$  enters the unemployed's utility function. This captures the notion of agents being social individuals who would like to spend their leisure time together, because there exists a complementarity in the consumption of free time. An agent needs e.g. other agents to play football with or have a chat in the pub. We only assume that the probability of finding an other individual with whom these activities can be shared increases with the overall leisure of all (unemployed) agents. Thus, leisure exerts a positive externality with  $\sigma$  capturing the strength of this effect.<sup>4</sup> The consequence of this externality is that individual's marginal costs of searching decrease with an increase in overall search intensity.

The probability of an individual to find a new job is given by:

$$q_i = s_i \theta^\alpha. \quad (5)$$

---

<sup>3</sup>The higher  $s_i$ , the lower is  $l_i$  and hence the higher is the marginal value of leisure (=marginal costs of searching). This is a standard assumption in the literature, see Pissarides (2000).

<sup>4</sup>We do not explicitly take into account leisure time of the employed since we assume their leisure time to be exogenously given, due to a fixed work contract. Since the time endowment of all agents is identical, we only have to consider the leisure time the unemployed have in excess of the employed.

Hence, the larger the individual search intensity (given aggregate search intensity in the economy) the higher will be the probability of finding a job.

When choosing optimal search intensity, the unemployed person has to take into account the following trade-off. A lower  $s_i$  increases the flow utility of staying unemployed, but also increases the expected duration of the unemployment spell. The latter effect is negative since the value of being employed is higher than the value of being unemployed.

The foc of this problem is given by the derivative of equation (4) with respect to  $s_i$  (note that the individual agent assumes all aggregate variables to be unaffected by his or her choice):

$$\begin{aligned} -\gamma l_i^{\gamma-1} l^\sigma + \theta^\alpha (W_i - U_i) &= 0, \\ \Leftrightarrow \gamma l_i^{\gamma-1} l^\sigma &= \theta^\alpha (W_i - U_i). \end{aligned} \tag{6}$$

In deriving this equation we took advantage of the fact that the value functions  $U_i$  and  $W_i$  depict values of being unemployed and employed respectively in the case of optimal behaviour of agents (see, e.g. Dixit and Pindyck (1996) or Shimer (2004) for a detailed argumentation). As such, these are *not* general functions of  $s_i$ . Thus, when deriving the condition for the choice of optimal search intensity, we do not have to take changes in the optimal state values into account.

The interpretation of the foc is straightforward. The lhs of the equation depicts the marginal cost of increasing search intensity which is the loss in flow utility due to less leisure time. The rhs on the other hand depicts the marginal value of higher search intensity which is equal to the increase in the probability of finding a job times the (optimal) net value of having a job.

The foc of individual's search behaviour depicts two important points: first, an increase in labour market tightness  $\theta$  will c.p. increase individual search intensity.<sup>5</sup> Secondly, the net value of having a job must exceed some threshold level so that the agents will start investing into search intensity. Throughout the paper we will assume that this condition holds and the economy is not shut down.

---

<sup>5</sup>This is a point Shimer (2004) focusses on. He argues that this is counterfactual in a business cycle context. Note, however, that this result is only of partial equilibrium nature. The aggregate relation between search intensity  $s$  and labour market tightness  $\theta$  will become clearer later on.

## 2.3 Firm Behaviour

Firms choose whether to invest into offering a vacant job slot. The Bellman equations for a vacant and a filled job slot are given by:

$$rV = -cp + m/v(J - V) \quad (7)$$

$$rJ = p - w_i + \lambda(V - J), \quad (8)$$

where  $p$  is the productivity of a worker,  $c$  denotes search costs and  $m/v$  is the probability of finding an adequate worker and turning the vacancy into a job. By free entry, the value of a vacancy must be zero. The matching function implies  $m/v = s\theta^{\alpha-1}$ . So the Bellman equations can be written as:

$$J = cps^{-1}\theta^{1-\alpha} \quad (9)$$

and

$$J = \frac{p - w_i}{r + \lambda} \quad (10)$$

Combining these two equations gives the job creation curve:

$$\frac{p - w_i}{r + \lambda} - cps^{-1}\theta^{1-\alpha} = 0 \quad (11)$$

## 2.4 Wage Determination

The last element of our description of the economy concerns the wage equation. We assume the wage to be bargained between a worker and a firm upon meeting. To determine the wage rate the implicitly assumed timing decision must be made explicit:

1. Stage: *Agents* choose the amount of search intensity they want to invest rationally anticipating the outcome of the bargain. *Firms* determine the number of vacancies they want to offer also anticipating the bargained wage.
2. Stage: Agents and Firms meet and bargain the wage.
3. Stage: Vacancies are filled and production starts.

Thus, in the wage bargain the amount of search intensity has already been invested by agents and is thus fix. The bargained wage will solve the following Nash product:

$$\Omega = (W_i - U)^\beta J^{1-\beta}, \quad (12)$$

where  $\beta$  denotes the bargaining power of the worker. Moreover, we took advantage of the fact that by the free entry condition the value of offering a vacancy must be zero. The foc for the bargaining wage  $w_i$  reads (using equations (3) and (10)):

$$\beta J(w_i) - (1 - \beta)(W(w_i) - U) = 0. \quad (13)$$

This simple structure of the wage setting rule is only due to the fact that the value of being employed is linear in the wage, i.e. that workers are risk-neutral. We can rewrite this equation and get a relation between the value of a filled job (which is given by the free entry condition) and the (optimal) net value of being employed:  $W(w_i) - U = \frac{\beta}{1-\beta}J(w_i)$ .

Using (13) we can also derive an explicit solution for the bargained wage. Plugging the expressions for  $W_i$  and  $J$  from the Bellman equations, (3) and (10), into the rent splitting rule we get:

$$\frac{w_i}{r + \lambda} - \frac{r}{r + \lambda}U_i = \beta \frac{p - w_i}{r + \lambda} + \beta \frac{w_i}{r + \lambda} - \frac{r}{r + \lambda}\beta U_i, \quad (14)$$

which may be simplified to

$$w_i = \beta p + (1 - \beta)rU_i. \quad (15)$$

By the foc of the Nash bargaining solution  $W_i - U$  is given by  $\frac{\beta}{1-\beta}J(w_i)$  so that we can simplify equation (4) to:

$$rU_i = b + l_i^\gamma l^\sigma + s_i \theta^\alpha \frac{\beta}{1 - \beta} J. \quad (16)$$

Plugging equation (9) into this equation gives

$$rU_i = b + l_i^\gamma l^\sigma + s_i \theta^\alpha \frac{\beta}{1 - \beta} c p s^{-1} \theta^{1-\alpha}. \quad (17)$$

Eventually, the wage equation reads:

$$w_i = \beta p + (1 - \beta)(b + l_i^\gamma l^\sigma) + \beta c p \frac{s_i}{s} \theta. \quad (18)$$

At the individual level, the effect of higher search intensity on the bargained wage is ambiguous. On the one hand, the individual value of being unemployed decreases with  $s_i$ , so that the firm can offer a lower wage. On the other hand, the agent has to be compensated for the higher search intensity, because the vacancy costs decrease. Moreover, the individual wage will decrease with an increase in aggregate search intensity in the economy.



Before turning to the determination of the general equilibrium in the economy, we will demonstrate how the endogeneity of the wage influences the decision making of agents concerning search intensity. Using equation (13) and (9), optimal search behaviour of agents is driven by the following foc:

$$\gamma l_i^{\gamma-1} l^\sigma = \frac{\beta}{1-\beta} c p s^{-1} \theta. \quad (19)$$

The right hand side of equation (19) depicts the marginal value of additional search that holds with an endogenous wage. The interpretation is straightforward:  $cp\theta s^{-1}$  is the expected search costs the firm would have to bear if it did not fill the vacancy with the worker just met. This is thus the cost the firm will save if it employs the worker. This is the matching rent. The bargained wage will be such that the net value of having a job will be a fraction of this matching rent. If the costs of filling a vacancy increased, the marginal gain of search for an unemployed worker would increase (since the net value of being employed would increase). Thus, their optimal search intensity would increase. The left hand side is as before marginal utility of leisure. This is unchanged in the general equilibrium.

The condition for optimal search behaviour of agents leads us to the following:

**Proposition 1** *Aggregate search intensity has countervailing effects on the optimal choice of an individual's search behaviour. For small  $s$ , individual search intensity will decrease with aggregate search intensity and vice versa.*

**Proof 1** *Take the differential of (19) to note:*

$$\begin{aligned} & (\gamma(1-\gamma)l_i^{\gamma-2}l^\sigma) ds_i + (-\sigma\gamma l_i^{\gamma-1}l^{\sigma-1}) ds = \left( -\frac{\beta}{1-\beta} cps^{-2}\theta \right) ds \\ \Leftrightarrow (1-\gamma)l_i^{-1} \frac{\beta}{1-\beta} cps^{-1}\theta ds_i - \sigma l^{-1} \frac{\beta}{1-\beta} cps^{-1}\theta ds &= \left( -\frac{\beta}{1-\beta} cps^{-2}\theta \right) ds \\ & \Leftrightarrow \frac{ds_i}{ds} = \frac{\sigma l^{-1} - s^{-1}}{(1-\gamma)l_i^{-1}} \gtrless 0. \end{aligned}$$

*The ambiguity of this expression is driven by the numerator. With  $l = T - s$  the following holds:*

$$\frac{ds_i}{ds} \begin{cases} > 0 & \text{iff } s > \frac{T}{1+\sigma} \\ < 0 & \text{iff } s < \frac{T}{1+\sigma} \end{cases}$$

■

An increase in aggregate search intensity has two countervailing effects on individual's choice. On the one hand, higher  $s$  will decrease the matching rent. In this situation firms will find it easier to fill vacancies. As a consequence there will be more firms entering the market with open job slots. But with this the value of a filled job slot must decrease.<sup>6</sup> Thus, the net value of having a job decreases which in turn discourages individuals to invest into search intensity. On the other hand, the marginal cost will decrease with aggregate search intensity. This is the impact of the leisure externality. With all other individuals searching intensively it is very unlikely to find someone to spend the leisure time with. The marginal value of leisure decreases, i.e. the marginal costs of search decrease and individuals are tempted to search more intensively. The latter effect will be the stronger, the larger  $s$  (at least for  $\sigma < 1$  which we assume). So for high values of  $s$  this effect will dominate the first one and agents will increase individual search intensity.

## 2.5 General Equilibrium

We will now derive and analyse the symmetric general equilibrium in the economy, i.e. a situation in which all agents and firms behave identical  $s_i = s$  and  $w_i = w$ . In the symmetric equilibrium the three equations derived in the previous sections (the first order condition of individual agents (19), the job creation curve of firms (11) and the wage curve (18)) solve the model for the three endogenous variables  $\theta$ ,  $s$  and  $w$ .

The wage is a function of average search intensity and of labour market tightness:  $w = w(\theta, s)$  with  $w_\theta > 0$  and  $w_s < 0$ , where the subscript denotes the partial derivative with respect to this variable. The economic intuition for these properties is straightforward. The higher search intensity of unemployed, the lower is their value of being unemployed, hence the firm only has to pay a low wage (in the Nash bargaining interpretation: "the outside option of the unemployed decreases"). A tighter labour market increases the wage that firms are willing to pay since search costs are high.

Using this wage equation, there are two relationships left which determine the equilibrium of the economy, namely the job creation curve and the optimal search behaviour of a representative individual:

$$(1 - \beta)(p - (b + l^{\gamma+\sigma})) - \beta c\theta - (r + \lambda)cps^{-1}\theta^{1-\alpha} = 0, \quad (20)$$

$$\gamma l^{\gamma+\sigma-1} = \frac{\beta}{1 - \beta} cps^{-1}\theta, \quad (21)$$

---

<sup>6</sup>Since the higher  $s$  makes it easier to fill a job slot, the implicit barrier to entry which protects incumbent firms decreases. As such, the value of a filled job slot must decrease.

where we used the expression for the bargained wage and the fact that in a symmetric equilibrium every agent will choose the same amount of search intensity, hence  $l_i = l$ .

In order to derive comparative static results, we have to calculate the slopes of these two equilibrium equations in the  $\theta$ - $s$ -space. Let us consider the job creation curve first. Totally differentiating equation (20) yields:

$$\begin{aligned} ((1 - \beta)(\gamma + \sigma)l^{\gamma+\sigma-1}) ds - \beta cd\theta - ((r + \lambda)cps^{-1}(1 - \alpha)\theta^{-\alpha}) d\theta \\ + ((r + \lambda)cps^{-2}\theta^{1-\alpha}) ds = 0 \end{aligned}$$

From this, the slope of the job creation curve is given by

$$\frac{d\theta}{ds} = \frac{(1 - \beta)(\gamma + \sigma)l^{\gamma+\sigma-1} + (r + \lambda)cps^{-2}\theta^{1-\alpha}}{\beta c + (r + \lambda)cps^{-1}(1 - \alpha)\theta^{-\alpha}} > 0. \quad (22)$$

Both the denominator and the numerator are positive, hence the job creation curve is positively sloped in the  $\theta$ - $s$ -space. Higher search intensity of unemployed agents makes it c.p. more profitable to offer vacant jobs, since the wage the firm has to pay decreases and the probability that the vacancy is filled increases. This is the effect also present in Diamond (1982) where more activity on one side of the market (in this case on the side of the unemployed searcher) induces activity on the other side.

Next, we turn to the slope of the curve that describes optimal search behaviour of agents. Totally differentiating (21) yields the following equation:

$$\begin{aligned} \frac{d\theta}{ds} &= \frac{(1 - \gamma - \sigma)\gamma l^{\gamma+\sigma-2} + \frac{\beta}{1-\beta} cps^{-2}\theta}{\frac{\beta}{1-\beta} cps^{-1}} \quad (23) \\ &\Leftrightarrow \frac{d\theta}{ds} = ((1 - \gamma - \sigma)l^{-1} + s^{-1})\theta. \end{aligned}$$

The slope of the curve depicting optimal search behaviour in  $\theta$ - $s$  space is ambiguous. This ambiguity is driven the fact that  $(1 - \gamma - \sigma)$  needs not to be positive (in the aggregate, marginal costs of search intensity need not to increase). In the case without leisure externality, the slope of the curve is strictly positive, i.e. an increase in  $\theta$  will increase search intensity, and as such resembles the results derived in Pissarides (2000).

However, we want to concentrate in the following on the (in our view) more interesting and relevant case in which a leisure externality exists. In the analysis we will focus on the case of  $\sigma + \gamma > 1$ . This implies that the leisure externality is strong enough implying that marginal utility of leisure will increase at the aggregate level. With this, we can state the following:

**Proposition 2** *If  $\sigma + \gamma > 1$ , the curve showing optimal (equilibrium) search intensity of agents will be hump shaped.*

**Proof 2** *Note that the slope of the curve is driven by:*

$$\begin{aligned} s^{-1} - (\gamma + \sigma - 1)l^{-1} \gtrless 0 &\Leftrightarrow s^{-1} \gtrless (\gamma + \sigma - 1)l^{-1} \Leftrightarrow \frac{l}{s} \gtrless (\gamma + \sigma - 1) \\ &\Leftrightarrow \frac{(T - s)}{s} \gtrless (\gamma + \sigma - 1) \Leftrightarrow \frac{T}{s} - 1 \gtrless (\gamma + \sigma - 1) \Leftrightarrow \frac{T}{\gamma + \sigma} \gtrless s \end{aligned}$$

For  $\bar{s} = \frac{T}{\gamma + \sigma} < T$ . This expression will hold true in equality. For  $\bar{s}$  the slope of the curve depicting optimal behaviour will be zero. For  $s < (>)\bar{s}$  the slope will be positive (negative). ■

The  $s$  associated with the positively and negatively sloped part of the curve showing agents optimal behaviour are in the choice set of agents  $s \in [0; T]$ . As such, both parts of the curve are relevant for the determination of the equilibrium.

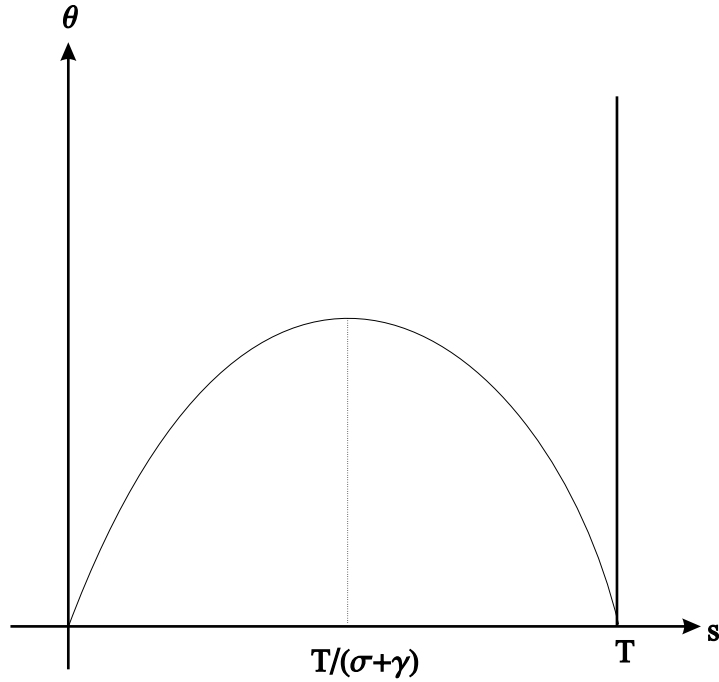


Figure 1: The Shape of Agents' optimal behaviour

Figure 1 depicts the curve showing aggregate optimal behaviour of agents implied by Proposition 2. The effect of a change in labour market tightness

$\theta$  on individuals behaviour depends on the level of aggregate search intensity. This is due to the fact that by the leisure externality the individual marginal costs of search intensity are a function of search intensity exerted by other agents.

Higher labour market tightness  $\theta$  will on impact induce agents to increase their level of search intensity independently of aggregate search intensity. This is because marginal gains of search will c.p. increase. However, and this is not internalised by the agents, individual decisions change aggregate behaviour, i.e. aggregate search behaviour changes. This in turn repercussions on the optimal behaviour of the individual. For low-values of  $s$ , individual search intensity will decrease with higher  $s$ . The second round effect, hence, dampens the impact effect of a higher  $\theta$ . But the impact effect will unambiguously dominate this second round effect. Search intensity will increase with higher  $\theta$ . This is the standard result also present in Pissarides (2000).

Explaining the negatively sloped part is not straightforward. Technically speaking, the curve depicts combinations of  $s$  and  $\theta$  where the marginal gain of search is equal to the marginal costs. An increase in  $\theta$  increases the marginal gain of search. As such,  $s$  must change in order to equate marginal costs and gains again. But this implies that for high levels of  $s$ , search intensity must decrease to close this gap. This reaction is due to the assumption of aggregate decreasing costs of search.

However, this is only a description of the curve, but not of the behaviour of agents in the case of  $s$  large enough. As before the increase in  $\theta$  will increase individual search behaviour  $s_i$  and thus,  $s$ . For high levels of  $s$ , however, this again will increase individual's search behaviour. Thus, any change in  $\theta$  will lead to an ever increasing search behaviour of agents. This process will not come to an end until all agents invested their entire time endowment into search. As such, all points depicted by the negatively sloped part of this curve are unstable. This is important for the characterisation of the equilibrium in the economy.

**Proposition 3** *If search intensity of agents is already very high, a change in labour market conditions will lead to a corner solution in which all unemployed agents will invest their entire time endowment into search.*

It is important to note that the corner solution is not the result of an explicit choice of agents. Every individual agent's utility function is well-behaved, i.e. at the individual level, agents face increasing costs of search activity. This usually would rule out corner solution. However, the leisure complementary between individual and aggregate leisure results in the above

described feedback effect (caused by aggregate decreasing costs of search intensity). This makes the economy end up in the corner situation.

After having characterised the behaviour of individual agents and firms, we can eventually turn to the equilibrium of the economy. The shape of the curves reflecting the search behaviour of agents and the job offer of firms (which depicts the supply of vacancies for given search intensities) imply that the economy is characterised by two-equilibria.

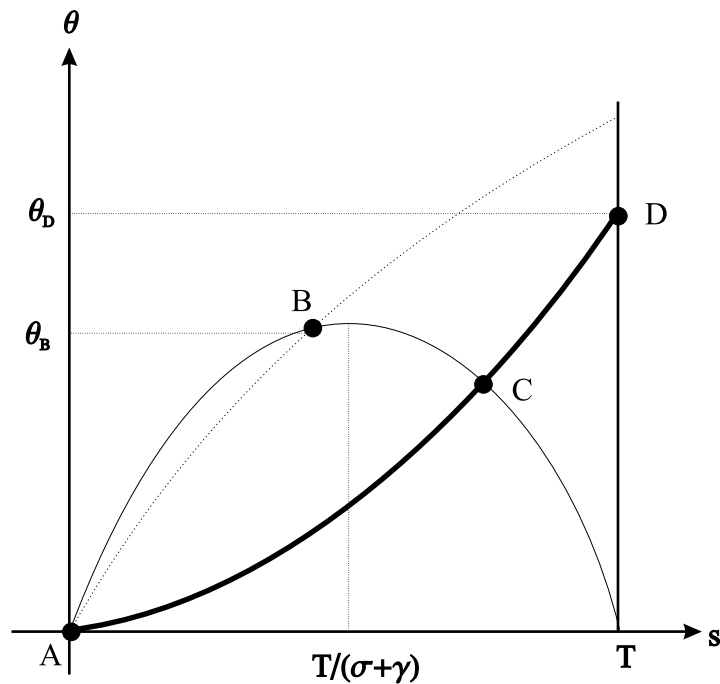


Figure 2: Equilibria in the Economy

Figure 2 depicts the equilibrium in the economy for different job offer constellations. The dashed line corresponds to a situation in which e.g. productivity in the economy is large. The thick line delineates a job offer curve for different parameter values. By the form of the two equilibrium forming equations (20) and (21) there exist at most two (real) solutions for the equilibrium.

Since both equations start at the origin, point  $A$  is always an equilibrium. This point reflects the "no-action" equilibrium. This equilibrium is basically driven by the interdependence of the actions of agents on both sides of the market. If e.g. firms expect agents not to invest into search intensity, they know that no vacancy will be filled. Thus, no vacancies will be offered. But with no vacancies being offered no agent will invest into search intensity. As

such, this equilibrium is the result of a coordination failure as described e.g. in Diamond (1982).

The "no-action" equilibrium, however, is not stable. A marginal increase in  $\theta$  will make an individual agent increase search intensity. This in turn makes all agents search more intensively until the new equilibrium is reached. Depending on the form of the job offer curve this could be a point such as  $B$ , which is stable or a point such as  $C$  which is not stable. If the economy is characterised by a job offer curve corresponding to the thick line in figure 2, the economy will end up in the corner solution  $D$ . This is due to the fact that (as described above) any change in labour market tightness  $\theta$  in a situation in which  $s > \frac{T}{\gamma+\sigma}$  will kick-off a self amplifying process of ever increasing search intensity until the corner is reached. In this situation firms will offer vacancies such that  $\theta = \theta_D$ .

The equilibrium values of aggregate search intensity and labour market tightness determine the equilibrium (steady state) unemployment rate in the economy. The change of the rate of unemployment in the economy is given by:

$$\dot{u} = \lambda(1 - u) - s\theta^\alpha u = 0.$$

In a steady state unemployment rate is thus given by:

$$u^* = \frac{\lambda}{\lambda + s\theta^\alpha},$$

which is only a function of  $s$  and  $\theta$ . Thus, the above derived equilibrium unambiguously determines the unemployment rate in the economy. In the above illustrated case the corner solution is associated with a lower rate of unemployment than the "standard" equilibrium, point  $B$ , since both search intensity and labour market tightness are larger. Note, however, that this does not hold necessarily for all parameter values.

Our model suggest that independently of whether social interaction exists or not, the economy is only characterised by one labour market equilibrium. This contrasts existing literature which states that social interaction will lead to multiple labour market outcomes. It was argued in the introduction that these two competing views of the world have important policy implications. Thus, it is important which of these views is reconcilable with the data. In the next section we use unemployment spell data for Germany to analyse whether regional labour markets are characterised by multiple equilibria or not.

## 3 Empirical Analysis

### 3.1 The Data

To employ the identification strategy described below, we use the so called Regionalfile of the IAB employment sample (IABS-R01) provided by the Bundesagentur für Arbeit (BA), Nuremberg. The IABS-R01 is a unique micro-level data set, including the employment history as well as the history of unemployment benefit receipt for two percent of all German employees subject to social insurance contributions for the period 1975 to 2001.

Two different sources of information are used for the creation of this data set. First, the dates concerning the employment history of individuals are generated using information provided by the social insurance institutions to whom employers yearly report the employment status of their employees, including daily dates on the beginning and/or ending of an employment spell. The data were amended by the periods of unemployment benefit receipt as supplied by the BA. It is important to notice that unemployment benefit payments do not correspond to social assistance since the latter are not provided by the BA. Thus, unemployed individuals receiving social assistance are not included in the sample.

The data set contains information on 1293819 individuals, including 181058 in East-Germany since 1992. Beside this huge amount of data, the IABS-R01 has two further advantages. Both the data on the individual employment status and the data on benefit receipt are provided in spells with daily exact dates on the beginning and ending of the spell. Beyond this, the data set features a regional clustering with 343 identifiable regions since 1980. The regions in the data set are entities on county-level and comparable to actual German counties. In order to provide a certain degree of anonymisation, the actual counties have been aggregated to regional entities with at least 100000 inhabitants yielding 343 regions in the data set. With this feature any individual employment or unemployment spell can be matched to a specific region.

Additionally the data set includes information on individual characteristics like age, sex, education, income while employed, occupation, etc. as well as information on the individual employment status (part-time/full-time, internship, apprenticeship, etc.) or sectoral affiliation while employed or before unemployment.



## 3.2 Empirical Strategy

### 3.2.1 General Idea

The purpose of the empirical strategy is to discriminate which impact social interaction have on the equilibrium structure of the economy. Is it true that regional labour markets are characterised by multiple or by a unique equilibrium? Thus we basically test the implication of the presented theoretical model.

In doing this we estimate and analyse the hazard for leaving unemployment, i.e. the probability of leaving unemployment conditional on being unemployed. If regional labour markets were characterised by a unique equilibrium, the hazard rates of unemployed agents should be identical between regions. Thus, we would only have to analyse the regional hazards rates and compare their distribution in order to discriminate between models. A drawback of this simple approach, however, is that it totally neglects the differences in hazard rates which are due to regional or individual heterogeneity, e.g. due to self-sorting effects. Consider that the population in region  $i$  is better educated than the one in region  $j$ . If we believe in education as an important determinant for the chances of getting a job, our simple approach would identify significant differences in the hazard rate between the regions. These differences, however, would be only due to composition effects. The same would be true for other differences in structural parameters between regions. Remember that social interaction models generating multiple equilibria that regional labour market performance differs although the structural characteristics are identical.

Thus, we apply an approach incorporating regional fixed effects. This results in regional hazard rates which control for individual and regional heterogeneity. We could interpret the regional hazard rates as the hazard rates identical individuals would face if all regions were homogenous. Analysing the differences between these hazard rates could give a hint which type of model is the superior one. If these regional differences were very small (large) we would conclude that when controlling for structural differences the regional labour markets perform identical (differently). This result would point to labour market being characterised by a unique (multiple) equilibrium (equilibria).

### 3.2.2 Estimation and further Procedure

#### Estimation

Since we are interested in an analysis of unemployment spells, we have to apply econometric methods of survival analysis which are appropriate to handle duration data. To measure the controlled individual unemployment durations mentioned above in the context of survival analysis, the so called baseline hazard rate provided by continuous parametric Proportional Hazard (PH) or Accelerated Failure Time (AFT) models can act in this way. In general, the baseline hazard rate corresponds to the situation where all covariates are equal to zero. Thus it can be interpreted as the hazard rate which is "common to all people" (see, e.g. Jenkins (2004), p.28) or in other words, the basic probability of an featureless individual for leaving unemployment in the next short instant of time, conditional on her actual unemployment duration. By estimating individual baseline hazards for each region, we get a proxy for these controlled individual unemployment durations and can investigate in a next step if significant variations in the individual baseline hazards over the regions exist. If the data actually features this property, we can assess this as an argument for multiplicity in regional labour markets.

To acquire estimates of the baseline hazard, a precise assumption concerning the distribution of the underlying survivor function has to be made. Therefore several PH and AFT models with different specifications for the survivor function will be estimated. In a second step we will choose among the models for further calculations.

With data on duration, one way to control for unobserved heterogeneities is to use a stratified estimation technique. Considering this, we get the following specification for the PH models

$$h_j(t|x_i) = h_{0j}(t)exp(x_i\beta), \quad (24)$$

where  $t$  is the duration of the unemployment spell,  $h_{0j}(t)$  is the baseline hazard for region  $j$ ,  $x_i$  is a vector of individual and regional covariates,  $\beta$  is a vector of parameters. Individual observations are indexed by  $i$ . For AFT models the specification below is estimated:

$$\ln t_i = x_i\beta + z, \quad (25)$$

where  $z$  is the error term.

In PH models the distributional form of the survivor function determines the shape of the baseline hazard rate  $h_{0j}$ . The idea of PH models is that the common baseline hazard is rescaled by the values of the covariates and the coefficients subject to a specific individual. Contrary, for AFT models

the distributional term of the error term  $z$  determines the regression model and with this the underlying distribution of the survivor function. The basic idea of AFT models is that the survival time of an individual is extended or shortened due to the values of the covariates and coefficients. Hence, in the context of AFT models, the term baseline hazard is not an explicit feature of the model. Nevertheless it corresponds to the estimated hazard rate with all covariates equal to zero.

For the estimations of the regional baseline hazard rates, we used 5 different models, assuming the following distributions for the survivor function: Exponential, Weibull, Gompertz, Log-Normal and Log-Logistic.

### Choosing among models

As already mentioned, we estimate several PH or AFT models, based on different assumptions concerning the distribution of the survivor function. With this a different duration dependence of the baseline hazard is generated in every model. Consequently the obtained results for the variation of the baseline hazards may vary between the several models. Thus, some words concerning model selection are in order.

As proposed by Klein and Moeschberger (1997), a popular way to choose among models is to use the Akaike information criterion (AIC) by Akaike (1974) or likelihood ratio tests. Unfortunately these ways of model selection fail with this analysis. The reason for this inappropriateness is the fact that we are explaining individual outcomes with variables measured at the regional level. This may lead to a violation of the assumption of uncorrelated errors and with this to biased standard errors as described in the regression context by Moulton (1990). Since in all estimates, inferences are based on robust standard errors, likelihood ratio tests or any other measures like the AIC based on log-likelihoods are not appropriate. Therefore the only way to evaluate the fit of the regressions and to choose among the several distributions and models is to rely on plots of Cox-Snell residuals as proposed by Klein and Moeschberger (1997). The idea behind this procedure is that if the correct model has been estimated, the Cox-Snell residuals will have a unit exponential distribution with a hazard ratio of 1.<sup>7</sup> To test whether the Cox-Snell residuals are unit exponentially distributed, plots of the integrated baseline hazard based on Cox-Snell residuals versus the Cox-Snell residuals could contain the desired information. If the posited model fits the data, the plot of the integrated baseline hazard based on Cox-Snell residuals versus the Cox-Snell residuals should yield a straight line through the origin with slope

---

<sup>7</sup>See Cox and Snell (1974), Collet (1994) and Klein and Moeschberger (1997).

equal to 1.

Beside the fit of a single model, we have to consider another topic in model selection. If the investigation of the Cox-Snell residuals suggest a PH model, we have to test the PH assumption of time-invariant regression coefficients. One of the major properties of PH models is the fact that the regression coefficients are assumed to be constant over time. If the data lead to time-varying coefficients in a PH specification, a different model would be more suitable to fit the data. To control or test for this assumption, we follow the procedure of a piecewise regression as proposed by Box-Steffensheimer and Jones (2004). By a piecewise regression we mean to estimate the relevant PH model for observations whose survival times fall above or below some predetermined values and to assess if the estimated covariate effects are consistent over time.

## Covariates and Data in Use

Individual unemployment durations from the years 2000 and 2001 will be used in the regressions. We decided to use spells from this period of the data set since it were the most current data we dispose of. By restricting the data set on two years, the problem of left-censoring could occur, since we may lose information concerning the individual unemployment history prior to 2000. To avoid this problem, we only focus on individuals which became unemployed on January 1st 2000 or later. After preparation the data set includes 61103 individuals with 65053 unemployment spells.

To control for individual and regional characteristics, we include several covariates in the regressions. Individual characteristics include the age of the observed individual at the beginning of her unemployment spell. Also dummy-variables for sex and educational attainment are embedded. In case of educational attainment, the data allows us to distinguish between four different categories: no professional training, secondary school and professional training, university-entrance diploma and university diploma (including degrees from universities of applied sciences).

Beside age and education also the former occupation of an individual may influence unemployment duration or the probability of leaving it. To control for this influence, dummy-variables indicating the economic sector in which the individual worked before she became unemployed are enclosed. Though the data set provides a very detailed partitioning concerning the economic sector, we decided to group the data in three economic sectors: agriculture, industry and goods and services. Due to missing data concerning the classification of the economic sector, we excluded 20 regions <sup>8</sup>, leaving

---

<sup>8</sup>The excluded regions are the cities of Braunschweig, Oldenburg, Remscheid, Solingen,

323 regions to analyze.

To measure existing regional heterogeneities we include several regional covariates. Beside the regional population, the number of people in each category of educational attainment are included. In case of the regional industry structure, the numbers of inhabitants working in each sector are used as covariates. The latter covariates serve as proxies for the economic structure of a region.

An important determinant for the success of a job search are the overall labour market conditions in a region. To control for these, two alternative measures are conceivable: the regional unemployment rate or the regional in-flows into unemployment. In accordance to Lalive and Stutzer (2004) we use the regional in-flows into unemployment to avoid any endogeneity problems which may be caused by using the unemployment rate. The regional in-flows as well as all the other regional covariates are measured for every year. Moreover, all covariates measure the overall situation of the individual or region at the beginning of each unemployment spell.

As suggested by Kiefer (1988), all covariates are measured as deviations from the mean. Additionally no constant term is included in the regressions. This allows us to interpret the estimated baseline hazard rates as the hazard rate of an average individual living in an average region. Beyond this, we defined reference categories for sex, industrial sector and secondary education.

### **Evaluation of the estimated baseline hazards and the Dip-Test**

After the regressions and model selection the baseline hazards for every region is calculated. We decided to calculate the region specific baseline hazards at different points in times. In detail, calculations have been made for unemployment durations of 30, 60, 90, 180, 270, 365, 455, 545, 635 and 730 days. At each of these points in time, we have to evaluate whether there are significant variations among the regional baseline hazards. Remember that a unique regional labour market equilibrium implies that the regional baseline hazards should be more or less identical. To formalise this point we argue that the distribution of regional baseline hazards should be characterized by a unimodal density function if the equilibrium is unique. On the other hand, if the data does not support the theoretical model, the baseline hazards should be significantly different over regions, leading to a multimodal

---

Offenbach am Main, Heidelberg, Regensburg, Erlangen and Fürth as well as the following counties: Enzkreis, Vogelsbergkreis, Tuttlingen, Mühl Dorf am Inn, Erlangen-Höchstadt, Fürth, Aichach-Friedberg, Neu-Ulm, Nürnberger Land, Biberach, Bördekreis and Ohre-Kreis.

density/distribution function. To control for such uni-/multimodal distribution of regional baseline hazard rates, we employ the so called Dip-test as proposed by Hartigan and Hartigan (1985).

In what follows, we only give the basic intuition of the Dip-test and refer for details to the original paper. The Dip-test is a non-parametric test for the unimodality of probability distributions. Consider some empirical distribution function, (which could be constructed by using observed values). Moreover, consider the set of all unimodal distribution functions. The latter are convex over  $(-\infty, a]$  and concave over  $[a, \infty)$ , where  $a$  is the mode of the distribution (which needs not be unique). Then a procedure is employed to derive a test statistic. This test statistic is the measure of the maximum distance between the empirical distribution function and the unimodal distribution function that a.) is nowhere greater than the empirical distribution function, and b.) minimizes this maximum distance between the two functions ( $\equiv$  Dip). If e.g. the empirical distribution function is an element of the set of all unimodal distribution functions, the distance/Dip is obviously zero. If the Dip is non-zero, this will imply that the empirical distribution departs from the "best-fitting" unimodal distribution which in turn is a hint that the empirical distribution could have more than one mode. Comparing the computed distance with distances computed under the null-distribution (which is the uniform distribution)<sup>9</sup> we can reject/accept the null-hypothesis that the distribution is unimodal.

### 3.3 Results

#### 3.3.1 Model selection and Regression Results

In a check of the Cox-Snell residual plots, the Gompertz and the Log-Logistic model seem to provide the best fit for the data.<sup>10</sup> We will concentrate on these models for the further investigations and look in a first step on the estimates of the regression coefficients. Table 1 reports the results and the z-values.

Table 1 around here

When it comes to interpret and/or compare the two models we have to keep in mind that the Gompertz model refers to the class of PH models while the Log-Logistic model is an AFT model. Recall, the coefficients of the Gompertz (PH) model describe the proportional influences of the covariates on

---

<sup>9</sup>Since this is the distribution function which is unimodal and in favor of multimodality.

<sup>10</sup>The plots are provided by the authors on request

the hazard rate, whereas the coefficients in the Log-Logistic (AFT) model contain information on the marginal influence of the covariates on the natural logarithm of the individual unemployment duration. Paying attention to this, we can observe that with one exception the two models exhibit equal results concerning the sign of the parameters.

Investigating the coefficients in detail, we find that in both models the dummies for educational attainment have the expected signs and are highly significant. An individual with university-entrance diploma or university diploma will leave unemployment faster than an individual with lower educational attainment. Also the proxy for individual occupation, the individual sector affiliation before unemployment, turns out to be highly significant. Finding a new job is more difficult for individuals who worked in the agricultural or in the goods and services sector, than for those working in the industrial sector.<sup>11</sup>

An interesting result yields the (highly significant) estimate of the impact of individual age for the hazard rate of leaving unemployment. In both regression an older individual tends to leave unemployment earlier. This is in a way counterintuitive since we expected younger individuals to have an earlier transition from unemployment to employment. One possible way to explain this result may be the problem of unemployment among young people. If we look at age-specific unemployment rates for Germany we find that for individuals of age 15 to 19, the unemployment rates are relatively small. Whereas in the second age cohort (20-24 years) unemployment rates are comparable to those for people of age 50 or older. The overall situation for young people after they finished e.g. apprenticeships or any other forms of professional training is therefore quite bad, whereas for people older than 24 the overall labour market situation turns out to be better. As a consequence, unemployment durations for individuals older than 24 should be shorter than for younger participants of the labour market. Thus the observed positive (negative) correlations between age and the hazard rate (survival time) may capture this effect.

For the impact of gender we find a positive (negative) influence on the hazard rate (unemployment duration). Thus, males face a higher probability of leaving unemployment and hence, face shorter unemployment spells. Economic intuition for this effect could be that in most cases males are the bread winners of a family and thus cannot be that picky when it comes to the decision of accepting or rejecting a job. This, effect should vanish when controlling for family status. Alas, the data does not provide us with this information.

---

<sup>11</sup>Recall we defined secondary education, male and industry as reference categories.

If we now turn to regional characteristics, we observe that most of them are insignificant. This is particularly true for the industry structure as well as for the number of inhabitants. We also obtain this result for the regional labour market situation as measured by the number of in-flows into unemployment. In both regressions a positive (negative) influence of the in-flows on the hazard rate (unemployment rate) is found. We could interpret this result as a characteristic of dynamic regional labour markets. In these labour markets the turn-over is very large, i.e. the probability of finding a job is very high.

Only three regional covariates have a statistically significant influence on the hazard rate (unemployment duration) in the regressions. Both the number of inhabitants with no professional training and the number of inhabitants with university diploma show the expected sign. The better the educational attainment in a region, the better individual labour market performance with respect to unemployment durations. Though we have a quite surprising sign on the individual level, on the regional level the age structure exerts a negative influence on the hazard rate/unemployment duration.

Beside individual and regional covariates, 323 regional dummies in the rescaling term, as well as 322 regional dummies and a constant in the estimation for the ancillary parameter of the assumed distribution were used. The coefficients for regional heterogeneity in the rescaling term of the hazard rate (unemployment duration) for the equation range from -76,63505 to 3,039078 for the Gompertz model and from -5,911557 to 72,48358 in the Log-Logistic case. In both regressions, the minimum/maximum value can be associated with the city of Berlin, stating that the chances of an individual to find a new job are extraordinary bad there. With the estimation for the ancillary parameter, coefficients range from 0,0004597 (-0,5093103) to 0,0079261 (0,1303587) in the Gompertz (Log-Logistic) model. While about two-thirds of the dummies in the rescaling term of the Gompertz model were significant, the estimates in the Log-Logistic model as well as for the ancillary parameter show a different picture. Only 25 of the coefficients in the Gompertz model and none in the Log-Logistic case were significant.

Summarising the regression results at this point of the analysis, we find that individual characteristics seem to be much more important for a fast transition from unemployment to employment than the economic conditions in the regional surrounding. Moreover, the small number of significant regional dummies in the estimation of the ancillary parameter may be a first indication that the baseline hazard of an individual is free of regional influences which is in accordance with the theoretical model presented above.



## Testing for the Proportional Hazard Assumption

Although the Gompertz model provided comprehensible estimations for the regression coefficients as well as the best fit in sense of the Cox-Snell residuals, it is still questionable whether the Gompertz model describes the data in an appropriate manner. This is due to the fact that models based on a Gompertz distribution belong to the class of PH models. As mentioned above, we control for the assumptions of time-invariant regression coefficients by estimating several piecewise regressions for different survival times.<sup>12</sup> Unfortunately the coefficients were not consistent over time which leads us to the insight that the Gompertz model is not an appropriate description of the underlying data. Therefore, the following calculations are based on the Log-Logistic model.

### 3.3.2 Baseline Hazards and the Dip-Test

#### Investigation of the Baseline Hazards

Since the Log-Logistic model belongs to the class of AFT models, we do not have an explicit baseline hazard as provided by PH models. But we can calculate it the same way, assuming all covariates to be zero. In general, the hazard rate for the Log-Logistic model is given by:

$$h(t) = \frac{-\partial \ln S(t)}{\partial t} = \frac{\lambda \gamma (\lambda t)^{\gamma-1}}{1 + (\lambda t)^\gamma} \quad (26)$$

with  $S(t) = \frac{1}{1+(\lambda t)^\gamma}$  as the survivor function for the Log-Logistic model. In the regression context,  $\lambda$  has been parameterized by the expression  $\exp(-x_i \beta)$ , with  $x_i$  as vector of individual and regional covariates and  $\beta$  as the vector of the estimated parameters. Assuming all  $x_i$  to be zero, we get the following expression for the baseline hazard:

$$h_0(t) = \frac{\gamma t^{\gamma-1}}{1 + t^\gamma} \quad (27)$$

Since we used a stratified estimation technique, we are able to calculate a baseline hazard for every region based on the results presented above, using:

$$h_{0j}(t) = \frac{(cons + d_j)t^{(cons+d_j)-1}}{1 + t^{(cons+d_j)}} \quad (28)$$

---

<sup>12</sup>Regressions have been made for the following periods: 0-30, 30-60, 60-90, 90-180, 180-270, 270-365, 365-455, 455-545, 545-635 and 635-730 days. The results of the regressions are available on request.

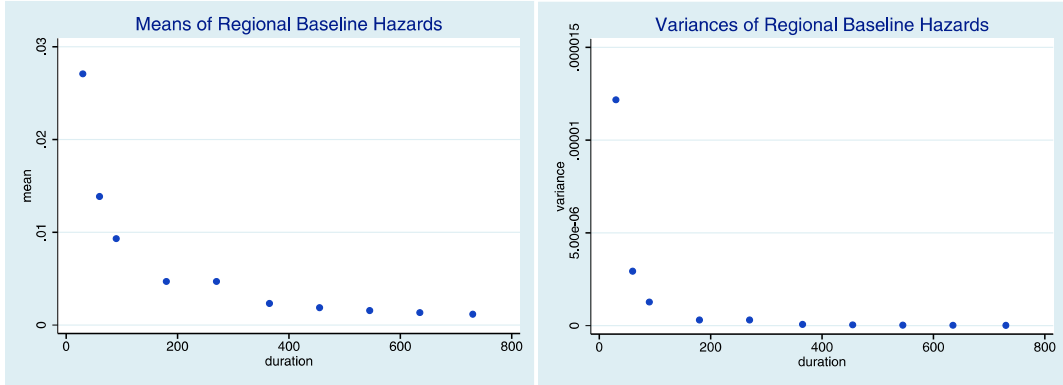


Figure 3: Means and Variances of the estimated Baseline Hazards

where  $cons$  is the estimated constant and  $d_j$  the dummy for region  $j$  in the regression for the ancillary parameter  $\gamma$ .

Using (28) we calculated regional baseline hazard for durations  $t$  of 30, 60, 90, 180, 270, 365, 455, 545, 635 and 730 days. The results of this are depicted in table 2.

Table 2 around here

Inspecting the means shows that regional baseline hazards tend to diminish over time. This suits well with general intuition that individual chances for leaving unemployment worsen with increasing unemployment durations, leading to the problem of long-term unemployment. Considering the resulting variances, the following points can be observed: First, the variances are very small, indicating only little differences between regional baseline hazards. This may be a further hint for the correctness of the theoretical model. Beyond this, variances also diminish over time indicating roughly equal baseline hazards with longer unemployment durations. In this respect, the general individual probability of a transition out of unemployment tends for longer unemployment spells to be independent from the location. Thus, the equality of regional baseline hazards seems to be stable over time. Figure 3 illustrates these results.

### Results from the Dip-Test

The variance of the distribution of regional hazard rates gives a first hint concerning the uniqueness of the equilibrium of regional labour markets. However, to analyse the distribution more thoroughly and to get a more

complete picture we use kernel density estimations to characterise the distribution (the density distributions for various duration are depicted in figure 4 in the appendix). Moreover, we apply a test for the uniqueness of regional hazard rates, the Dip-test.

As already explained, the idea for applying this test is the notion that if regional labour markets are characterised by a unique equilibrium, the regional baseline hazards should be (randomly) distributed around this equilibrium. This in turn implies that the density distribution of regional hazard rates should be unimodal. Thus the null-hypothesis (= unimodal distribution) of the Dip-test should not be rejected if this conjecture is correct. In what follows, we present the results of the Dip-test for the particular unemployment durations. It turns out that for all distributions of regional hazard rates the null-hypothesis of unimodality cannot be rejected (see table 3).<sup>13</sup> Consequently we have to conclude that with this identification strategy the data support the theoretical model. Therefore, the economy seem to be characterized by one equilibrium on the labour market, suggesting that social interactions among individuals have no effect on regional unemployment disparities.

Table 3 around here

Although the results of the Dip-test are very much in favour of the result of our theoretical model, one has to think about the robustness and the reliability of the test for identifying multiple equilibria.

Firstly, note that the Dip-test by its construction is very conservative, i.e. biased in favour of the null-hypothesis. As Cheng and Hall (1997) state, the use of the uniform distribution as a benchmark for the worst uni-modal distribution is responsible for this distortion. One way to handle this problem would be to rely on other tests for uni-/multimodality of distributions. Two popular choices in this area are the excess mass test by Müller and Sawitzki (1991) or the bandwidth test as suggested by Silverman (1981). However, according to Cheng and Hall (1997), these may also lead to considerable distortions of the results and are therefore not appropriate. Moreover, the Dip-test may be misleading in the case in which regional labour markets are characterised by multiple equilibria, but the difference between these equilibria is very small. Or in the case in which the majority of regions is determined by one equilibrium and only some regions are from another one.

---

<sup>13</sup>The test-statistic as well as plots of the densities are included in the appendix. All calculation concerning the Dip-test have been done with R 1.8.1 while for the survival analysis Intercooled Stata 8.2 has been used.

Also in this case the Dip-test would indicate unimodality. This is again due to the fact that the test is not sensitive enough.

Thus, although the test generates interesting results which are in favour of the notion that regional labour markets are really characterised by a unique equilibrium, we cannot unambiguously reject the hypothesis of multiple equilibria driving regional unemployment disparities.

## 4 Summary and Conclusion

Labour market models which incorporate preference based social interaction are often characterised by multiple equilibria. This is due to the fact that individual behaviour is influenced by the behaviour of the surrounding agents, which yields an externality. The consequence of this interdependence might be coordination failures. Thus, an identical agent would behave differently depending on the environment she lives in. Models with social interaction potentially could explain the disparities in regional unemployment rates.

In this paper, however, we present a Pissarides (2000) type search model in which we incorporate social interaction. Our notion of social interaction is modelled as a leisure externality. Unemployed agents have to invest time into the search process of finding a new job. The value of this time investment depends on aggregate time investment since agents are social individuals who would like to spend their time together. Despite this interaction, the economy is characterised by a unique (stable) labour market equilibrium. This uniqueness result is because of the existence of a "no action" equilibrium, i.e. if firms expect that agents do not invest time they will offer no jobs. If on the other hand agents expect firms not to offer jobs, they will not invest time into the search process.

If social interaction leads to multiple equilibria or not has important policy implications since the existence of multiple equilibria will broaden the scope for public regional policies. This begs the question whether we observe multiple equilibria in real world labour markets. Using a micro-level data set on the duration of unemployment in Germany we analyse whether regions are characterised by multiple equilibria. To do this we apply survival analysis to estimate the individual hazard rate of leaving unemployment controlling for structural individual and regional variables.

We find that the overwhelming part of differences in the hazard rate between individuals is explained by structural individual characteristics. Structural regional heterogeneity has surprisingly little effect on duration of unemployment. From this we conclude that for leaving unemployment it does not matter where you are, but who you are. Moreover, we find that the

hazard rate controlling for structural variables basically does not differ between regions. From this we conclude that the regional disparities in the unemployment rate can not be explained by different social interaction.

Social interaction provides us with a better understanding of the functioning of labour markets,<sup>14</sup> the data, however, does not support the view of multiple equilibria in regional unemployment rates as sometimes suggested in the literature, see e.g. Hedström et al. (2003).

## References

- [1] Akaike, Hirotugu (1974), A New Look at Statistical Model Identification, IEEE Transaction and Automatic Control, AC-19, pp 716–723.
- [2] Box-Steffensmeier, Janet and Bradford S. Jones (2004), Event History Modelling, The Cambridge University Press.
- [3] Cheng, Ming-Yen and Peter Hall (1998), Calibrating the Excess Mass and Dip Test of Modality, Journal of the Royal Statistical Society, 60, pp 579–589.
- [4] Collett, Dave (1994), Modelling Survival Data in Medical Research, London: Chapman and Hall.
- [5] Cox, D.R. and E.J. Snell (1968), A General Definition of Residuals (with discussion), Journal of the Royal Statistical Society B, 30, pp 248–275.
- [6] Dixit, Avinash K. and Robert S. Pindyck (1994), Investment under Uncertainty, Princeton University Press.
- [7] Elhorst, Paul J. (2003), The Mystery of Regional Unemployment Differentials: Theoretical and Empirical Explanations, Journal of Economic Surveys, 17, pp 709–749.
- [8] Glaeser, Edward L. and Jose Scheinkman (2000), Non-Market Interactions, NBER Working Papers: 8053.
- [9] Hartigan, John A. and P. M. Hartigan (1985), The Dip Test of Unimodality, The Annals of Statistics, 13, pp 70–84.
- [10] Hartigan, P.M. (1985), Computation of the Dip Statistic for Unimodality, Applied Statistics, 3, pp 320–325.

---

<sup>14</sup>Remember that the presented model can lead to corner solution implying different comparative statics behaviour than the standard model without leisure complementary.

- [11] Hedström, Peter; Ann-Sofie Kolm and Yvonne Åberg (2003), Social Interactions and Unemployment, IFAU Working Paper Nr. 15.
- [12] Hosios, Arthur J. (1990), On the efficiency of matching and related models of search and unemployment, *The Review of Economic Studies*, 57, pp 279–298.
- [13] Jenkins, Stephen P. (2004), Survival Analysis, Unpublished manuscript, Institute for Social and Economic Research, University of Essex.
- [14] Kiefer, Nicholas M. (1988), Economic Duration Data and Hazard Functions, *Journal of Economic Literature*, 26, pp 646–679.
- [15] Klein, J.P. and Melvin L. Moeschberger (1997), *Survival Analysis: Techniques for Censored and Truncated Data*, Springer-Verlag.
- [16] Lalive, Rafael and Alois Stutzer, The Role of Social Work Norms in Job Searching and Subjective Well-Being, *Journal of the European Economic Association*, 2(4), pp 696–719.
- [17] Manski, Charles F. (2000), Economic Analysis of Social Interactions, *Journal of Economic Perspectives*, 14, pp 115–136.
- [18] Montgomery, James D. (1991), Social networks and labour market outcome: toward an economic analysis, *American Economic Review*, 81, pp 1408–1418.
- [19] Moulton, Brent R. (1990), An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables in Micro Units, *Review of Economics and Statistics*, 72, pp 334–338.
- [20] Müller, Dietrich Werner and Günther Sawitzki (1991), Excess Mass Estimates and Tests for Multimodality, *Journal of the American Statistical Association*, 86, pp 738–746.
- [21] Nickell, Stephen and Richard Layard (1999), Labor Market Institutions and Economic Performance, *Handbook of labor economics*, 3C, pp 3029–84, Elsevier.
- [22] Topa, Giorgio (2001), Social Interactions, Local Spillovers and Unemployment, *The Review of Economic Studies*, 68, pp 261–295.
- [23] Topa, Giorgio and Timothy G. Conley (2002), Socio-Economic Distance and Spatial Patterns in Unemployment, *Journal of Applied Econometrics*, 17, pp 303–327.

- [24] Overman, Henry G. and Diego Puga (2002), Unemployment clusters across Europe's regions and countries, *Economic Policy*, pp 117–147.
- [25] Pissarides, Christopher A. (2000), *Equilibrium Unemployment Theory*, 2nd Edt., The MIT Press.
- [26] Selod, Harris and Yves Zenou, (2001), Social Interactions, Ethnic Minorities and Urban Unemployment, *Annales d'Economie et de Statistique*, 63-64, pp 183–214.
- [27] Shimer, Robert, (2004), Search Intensity, mimeo, University of Chicago.
- [28] Silverman, Bernard W. (1981), Using Kernel Density Estimates to Investigate Multimodality, *Journal of the Royal Statistical Society*, B 43, pp 97–99.

## **5 The Appendix**

Distribution of the survivor function Model	Gompertz PH		Log-Logistic AFT	
Variable	coefficient	z-value	coefficient	z-value
<i>Individual characteristics</i>				
age	0,0277937	37,42	-0,0343365	-36,5
male	0,1589121	10,08	-0,1872675	-9,47
no professional training	-0,2091514	-14,17	0,242458	12,25
university entrance-diploma	0,0939743	3,47	-0,1318498	-4,11
university diploma	0,2015893	8,34	-0,2347597	-7,9
agricultural	-0,1009957	-3,14	0,1206943	3,17
goods and services	-0,1232007	-8,57	0,1384611	7,6
<i>Regional characteristics</i>				
average age	-1,213748	-14,74	1,28988	13,52
Number of inhabitants with/working in				
no professional training	-0,0027907	-2,39	0,0040734	2,85
secondary education	0,001186	0,81	-0,0022332	-1,3
university-entrance diploma	-0,004011	-1,42	0,0042267	1,27
university diploma	0,0061929	2,66	-0,0074307	-2,77
agriculture	0,0016737	0,89	-0,0010273	-0,51
industry	0,0005114	0,34	0,0004098	0,22
goods and services	0,0007112	0,55	-0,000292	-0,18
Number of in-flows into unemployment	0,0015777	1,48	-0,0018946	-1,53
Number of inhabitants	-0,003771	-2,45	0,0035659	1,73
Number of regional dummies		323		323
min	-76,63505	-6,38	-5,911557	-2,37
max	3,039078	2,02	72,48356	3,99
Number of significant (>90%) regional dummies		211		99
<i>Ancillary parameter</i>				
Constant	-0,0073103	-1,78	0,0224276	0,07
Number of regional dummies		322		322
min	0,0004597	0,11	-0,5093103	-1,52
max	0,0079261	1,92	0,1303587	0,39
Number of significant (>90%) regional dummies		25		0
Number of individuals		61103		61103
Number of unemployment spells		65053		65053
Number of strata		323		323
Number of regions		323		323

Table 1: Regression Results



<b>Duration</b>	<b>Mean</b>	<b>Variance</b>	<b>Min</b>	<b>Max</b>
30	0,0270753	0,0000121745	0,0182302	0,0381113
60	0,0138534	0,0000029357	0,0094769	0,0192547
90	0,009315	0,0000012764	0,0064238	0,0128772
180	0,0047015	0,0000003087	0,0032793	0,0064574
270	0,0047015	0,0000003087	0,0032793	0,0064574
365	0,0023311	0,0000000732	0,00164	0,0031887
455	0,0018721	0,0000000468	0,0013199	0,0025586
545	0,0015642	0,0000000325	0,0011046	0,0021364
635	0,0013433	0,0000000238	0,0009498	0,0018338
730	0,001169	0,0000000180	0,0008274	0,0015953

Table 2: Descriptive Statistics for the Regional Baseline Hazards (n=322)

<b>Duration</b>	<b>Dip</b>	<b>Unimodality (Significance &gt; 90%)</b>
30	0,01142174	yes
90	0,01141233	yes
90	0,01141666	yes
180	0,01141405	yes
270	0,01141405	yes
365	0,01137247	yes
455	0,01138176	yes
545	0,01139106	yes
635	0,01138868	yes
730	0,01142376	yes

Table 3: Results of the Dip-test for Unimodality. The critical values for determining the level of significance have been calculated for  $n = 322$  and using the methodology of Hartigan (1985). The results are available from the authors on request.

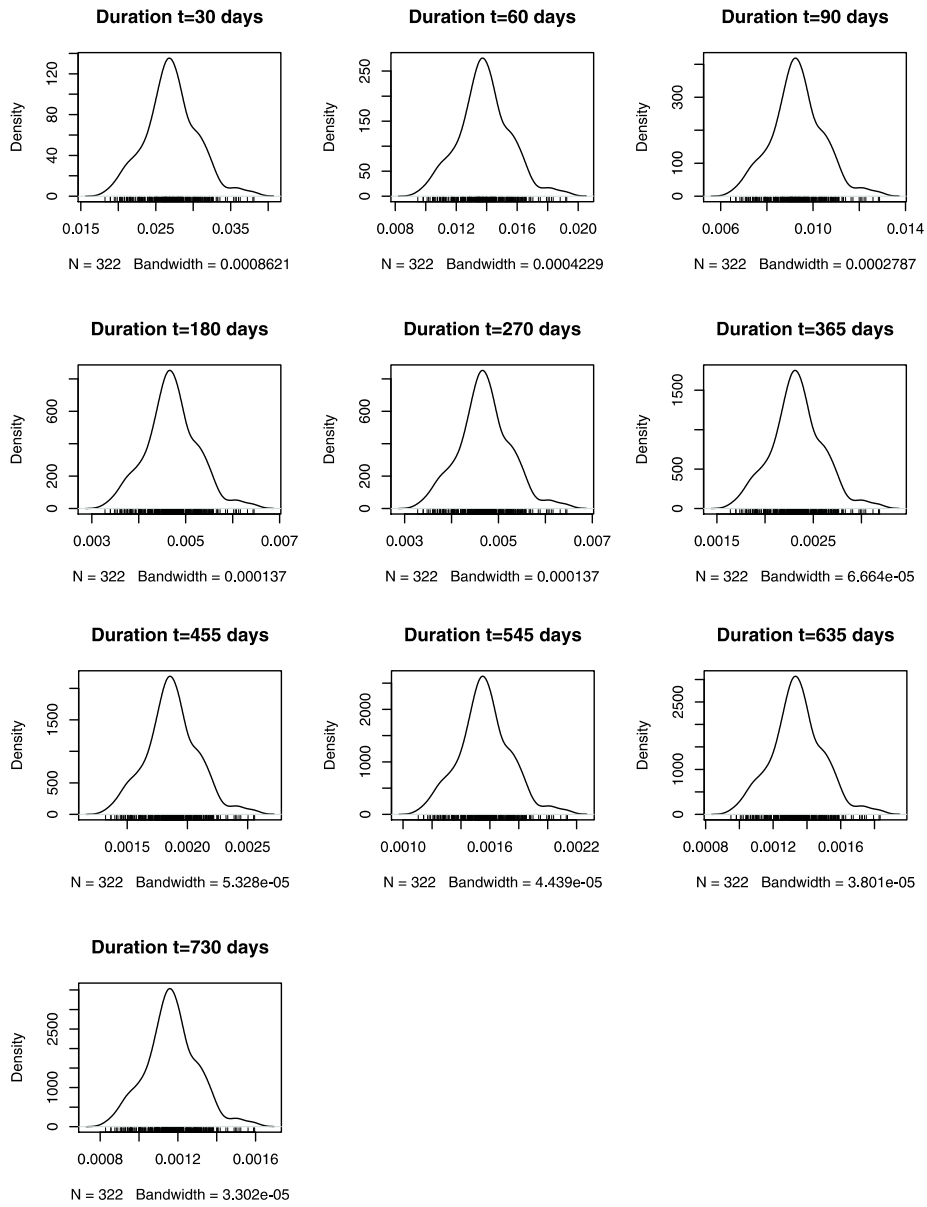


Figure 4: Kernel Density Estimates for the Baseline Hazard