

Labour market rigidities, unemployment and growth*

Fabiano Schivardi

January 2001

Abstract

This paper offers a critical survey of the literature on labour market rigidities, developing a basic theoretical framework that permits unified treatment of the various issues involved. First, it is shown how the simple equation linking rigidity with unemployment finds little theoretical or empirical support. At the same time, the degree of rigidity influences the dynamics, composition and duration of unemployment, with major consequences for the labour market, unemployment and the efficiency of the economy as a whole. We then analyze the interaction between employment protection legislation and the determination of wages; the effects of setting uniform wages in regions with different levels of productivity and immobile labour; and the consequences of the inefficient allocation of labour induced by a rigid system for capital accumulation and growth. The analysis shows that these factors can play an important role in understanding the disappointing performance of the labour market over the last three decades, independently of the direct effects of firing restrictions.

*This is the english version of a paper appeared in italian in the "Giornale degli Economisti e Annali di Economia", April 2000. Please also refer the reader interested in this version to <http://digilander.iol.it/fschivardi/>. I would like to thank Giuseppe Bertola, Tito Boeri, Andrea Brandolini, Rita Cappariello, Salvatore Chiri, Piero Cipollone, Francesco Lippi, Marco Magnani, Patrizio Pagano, Sandro Trento, an anonymous referee and the participants in a study group of the Bank of Italy Research Department for helpful discussions and comments. All errors are my responsibility. The opinions expressed in the paper do not necessarily reflect those of the Bank of Italy. E-mail: schiva@stanfordalumni.org

1 Introduction

The disappointing employment performance of the European economy during the 1990s, with unemployment rates on the order of 10 per cent, has sparked a lively debate on the efficiency of the institutions governing the operation of the labour market. The rigidity of this market is often cited as the principal cause of high unemployment and the economy's poor ability to create jobs. This thesis is indirectly supported by the performance of the US economy, which is often held up as an example of a flexible system, one that over the last decade has recorded high growth rates and steadily declining unemployment, which is now close to 4 per cent. This has been mirrored by a continuous increase in employment, which rose by about 18 million between 1990 and 1998.

This paper offers a critical synthesis of the debate on labour market rigidities, with an emphasis on its theoretical aspects. The main conclusion is that the theoretical analysis is more nuanced than the positions advanced in the policy debate. In particular, flexibility can have important indirect effects, while the simple equation linking rigidity with unemployment finds little theoretical or empirical support.

The first part of the work adopts a narrow, precise definition of rigidity, which is considered in terms of restrictions on changes in a firm's work force. Section 2 examines partial economic equilibrium models with given wages, which are the best developed part of the theory. The results that emerge from this class of models are that firing restrictions:

1. Do not have first-order effects on the average level of employment;
2. Constrain cyclical changes in employment, reducing flows into and out of unemployment and increasing its average duration;
3. Adversely affect the allocation of resources and reduce corporate profitability.

The empirical evidence provided by the OECD's Employment Outlook, which is briefly reviewed later, would appear to support the model predictions. It also emphasizes the indirect effects of rigidity. First, the duration of unemployment lengthens as the rigidity of the labour market increases, which can influence the equilibrium unemployment rate, since the long-term unemployed tend not to exert downward wage pressure that could help them re-enter employment. In addition, empirical analysis shows that employment

protection legislation could have different effects on the various components of the labour force, for example by favouring prime-age men over women and, above all, young people. The rigidity of the labour market could act as a barrier to entry with adverse consequences on youth employment, an hypothesis that merits further investigation.

Section 3 examines how the institutions regulating the labour market interact with labour supply and demand in determining employment and wages. The first result is that high rigidity tends to be accompanied by a high rate of real wage growth, since firing restrictions increase the bargaining power of the employed, protecting them from the downward wage pressure exerted by the jobless. It is also shown that wage compression, linked with prohibitions on paying different wages to workers with similar duties, is necessary to ensure that firing restrictions are effective, such that systems with little flexibility in quantities (labour) also tend to have inflexible prices (wages).

Section 4 illustrates the effects of wage compression at the local level, i.e. the imposition of uniform wages on regions with different productivity levels. It is shown that with low labour mobility, uniform national wages have an adverse impact on employment in the least productive regions, harming firms in those areas and workers in high productivity regions. Section 5 considers the “dynamic” effects of rigidity, focusing on the process of capital formation and growth. It is demonstrated that the lower profitability of firms in a rigid system can reduce the incentive to invest and thus reduce the rates of capital formation and balanced growth.

The main message that emerges from the survey is the importance of the indirect and selective effects of rigidity and of the need to broaden the analysis to other forms of rigidity. This paper is intended to be a starting point for improving our understanding of the relative importance of these issues and their degree of complementarity. This thesis is the focus of the concluding section, which examines the literature on practical strategies for reforming the labour market.

2 Firing restrictions and employment

This section analyses the role of a specific element of the conception of flexibility: firing restrictions. We conduct a partial equilibrium analysis, with fixed wages and exogenous demand (or productivity). In this setting, we examine the labour demand of a representative firm under two different

hypotheses on the adjustment costs of modifying labour.

Let there be two countries, {f, r}, one allowing total flexibility in the adjustment of the labour force (f), the other totally rigid, which we assume means that it is impossible to modify the labour stock. The two countries are otherwise identical. Each has one firm with the same production function, with labour as the only input:

$$y = A_i \log L, \quad i = \{h, l\} \quad (1)$$

where A_i is a demand (or technology) shock that can have one of two values $A_h > A_l$ with probability p and $1-p$ respectively.¹ The real wage is given at w . In country f , employment is decided after the occurrence of shock A_i , it is immediate to calculate the level of employment in the two states for a profit-maximizing firm:

$$L^f(i) = A_i/w, \quad i = \{h, l\}. \quad (2)$$

In country r the problem is only slightly more complicated, since employment must be decided before the shocks occur and cannot be changed ex post. In this case, the firm chooses employment so as to maximise expected profits:

$$\text{Max}_L E\{A_i \log L - wL\} = \text{Max}_L \{[pA_h + (1-p)A_l] \log L - wL\} \quad (3)$$

from which we obtain a level of employment equal to

$$L^r = \frac{pA_h + (1-p)A_l}{w} \quad (4)$$

independently of the state of the economy. The solution therefore produces an intermediate level of employment with respect to the two optimal values under perfect flexibility: $L^f(l) < L^r < L^f(h)$, since $\forall A_l < A_h$, $p \in (0, 1)$:

$$\frac{A_l}{w} < \frac{pA_h + (1-p)A_l}{w} < \frac{A_h}{w} \quad (5)$$

For simplicity, we assume that the labour force in the two countries is equal to A_h/w , so that country f achieves full employment in high productivity periods. The unemployment rate is equal to equazione. We obtain the following conclusions:

¹In the case of a demand shock, A is the price at which the firm sells the product, with a technology shock, a productivity factor for a given price.

- **Average employment** (and hence unemployment) is the same in the two countries:

$$\bar{L}^r = \bar{L}^f = \frac{pA_h + (1-p)A_l}{w} \quad (6)$$

Restrictions on dismissals have no impact on the average level of unemployment in country f, since lower employment in periods of high productivity //??// is offset by higher employment in periods of low productivity. In other words, firing costs symmetrically dampen changes in employment in both directions, leaving the average unchanged.

- The **variability of employment** (and hence unemployment) is higher in country f: the variance of the unemployment rate is:

$$\begin{cases} \text{var}(u^r) = 0 \\ \text{var}(u^f) = p(1-p)\left(\frac{A_h - A_l}{A_h}\right)^2 \end{cases} \quad (7)$$

In other words, the cyclical sensitivity of employment is stronger in the flexible country, increasing in recessions and falling in expansions; however, the net effect over the cycle is nil.

- The **duration of unemployment** is longer in country r: while the unemployed cannot find work in the latter, given that turnover is zero, the unemployed in country f find jobs with a probability of 1 if the state of productivity moves from A_l to A_h , which occurs with probability p , such that the average duration of unemployment is equal to $1/p$.
- **Average output and profits** are lower in country r. The demonstration of this result is given in the Appendix. Intuitively, while the average level of employment in the countries is the same, the flexible economy can allocate labour more efficiently, producing more during high-productivity periods. As for profits, since firms face an additional constraint in the rigid economy, their level of profits cannot be higher than that in the flexible economy.²

These are the principal conclusions in the literature regarding the effects of adjustment costs in a partial equilibrium context, obtained in a model

²Note that this also holds in a cross-section setting, where productivity differs across firms in a given period. For example, we can assume that p represents the share of high-productivity firms in each period. In this case, an efficient allocation of resources requires higher employment in firms with higher productivity, while in the rigid system, where adjustments are not possible, all firms have the same size labour force, regardless of their productivity.

with an extreme assumption about the size of such costs. The most general reference model has been developed by Bentolila and Bertola (1990), who study the optimal labour force adjustment policy for a firm facing demand that evolves in a Brownian motion process in continuous time, with linear firing and hiring costs. They conclude that adjustment costs have a significant impact in terms of turnover and the variability of unemployment over the cycle, but only second-order effects (and, in parameterizations, generally positive) on the average level of employment. Compared with perfect flexibility, in the presence of adjustment costs firms reduce employment less after negative shocks and increase it less after positive shocks, so that the average is less pronounced. The authors conclude that constraints on dismissals can explain the poor sensitivity of European employment to recoveries, but in themselves have little explanatory power with regard to the high level of average unemployment.

The results appear robust to various modifications to the theoretical specification. Abandoning the exogenous wage hypothesis, Blanchard (Blanchard 1998; Blanchard and Portugal 1998) propose a matching scheme in which wages are formed through bargaining between employed workers and firms.³ The analysis focuses on two variables: the duration of employment and the turnover rate. For employed workers, the relevant variable in the bargaining process is the duration of unemployment, i.e. how long they would be unemployed if they were to lose their job. An increase in duration prompts greater wage moderation. Since an increase in the cost of firing lowers the turnover rate and increases the duration of unemployment,⁴ it has an ambiguous effect on the equilibrium unemployment rate, which is given by definition as the product of duration and turnover. On the other hand, an increase in the duration of unemployment and a reduction in turnover radically modify the nature of unemployment. Blanchard and Portugal (1998, p. 3) summarize their empirical analysis of unemployment in the United States and Portugal as follows: “...if the image of the US unemployment is that of

³The literature on matching (Pissarides 1985; Mortensen and Pissarides 1994) describes the labour market as the place where the unemployed are matched with vacancies, assuming that this requires resources and time. The outcome of a match is a surplus that is split between the worker and the employer. In this context, the approach analyses the reaction of the process of job creation and destruction to various shocks, and what role is played by the rules for dividing the surplus in determining this reaction.

⁴The latter phenomenon is attributed to two factors: a) given that on average workers are less productive, wages must decline; as a result, duration increases to moderate wage demands; b) adjustment costs increase the bargaining power of workers, and once again longer unemployment is needed to prevent unsustainable wage growth.

a way station between jobs, the image of the Portuguese unemployment is that of a stagnant pool, with low flows in and out, and long unemployment duration” .

Garibaldi (1998) reaches similar conclusions in a matching model modified by the assumption that workers can only be fired after firms have received authorization to do so (for example, a court judgement), which is granted stochastically. Ljungqvist (1997) uses computational techniques in a general equilibrium model to assess the effects of firing limitations in three different classes of employment models, i.e. search models⁵ (Stigler 1961), matching models (Pissarides 1985) and real business cycle models⁶ (Rogerson 1988; Hansen 1985). For the first two classes, the effects of firing costs on average unemployment are ambiguous, generally tending to reduce the level, since lower turnover reduces frictional unemployment. The matching model produces a different result when firing restrictions increase the bargaining power of workers, in which case a higher level of unemployment is needed to re-establish firm profitability, reducing the average time needed to fill a vacancy. For the third class of model, firing costs tend to reduce employment, in line with results of the more general model of Hopenhayn and Rogerson (1993). However, the underlying assumptions are not entirely convincing: it is assumed that the tax on dismissals is returned to workers in the form of a lump-sum transfer, thus adversely impacting the labour supply. Alvarez and Veracierto (1998) extend the analysis of Hopenhayn and Rogerson (1993) by introducing search frictions in the reallocation process and incomplete insurance markets. Numerical solutions to the model show that in this case the introduction of firing costs reduces average unemployment, lowering frictional unemployment and increasing unemployed workers’ effort in the hunt for a job, since the job stability ensured by firing costs increases the desirability of employment itself.

⁵In search models, an unemployed person receives a job offer in each period at a given wage, extracted from a known probability distribution. The worker therefore must decide whether to accept the wage, which then remains fixed over time, or to reject the offer and continue the job search in the hope of receiving a better offer in the future. Unlike matching models, only the labour supply side is modelled. See Sargent (1987) for an introduction to the literature.

⁶In the models developed by Rogerson (1988) and Hansen (1985), it is assumed that the number of hours a person can work is indivisible owing, for example, to administrative restrictions on the duration of the working day. As a result, rather than choosing the quantity of work at a given wage, agents choose the probability of working at that wage. This approach means that while the individual labour supply function may be inelastic with respect to the wage, the aggregate function will be infinitely elastic.

2.1 The empirical evidence

This paper does not attempt a complete survey and critical assessment of the empirical evidence. The empirical analysis of the effects of labour market rigidity faces considerable methodological difficulties, due among other things to the difficulty of measuring the rigidity of a system, of establishing the direction of the causality between rigidity and unemployment and of taking account of the lags between institutional changes and their impact on the labour market. These problems counsel caution in interpreting the evidence. This section has the more modest aim of highlighting a number of “stylized facts” that emerge from the literature, in order to conduct a preliminary evaluation of the models surveyed in the previous section. The OECD’s Employment Outlook 1999 (OECD 1999) contains a survey of the empirical literature and a series of new results that the interested reader is invited to consult. Here we briefly summarize the main conclusions of the OECD study, which constructed a number of labour market rigidity indicators for the OECD countries, which were then used in cross-section regressions to study the effects on the labour market.

In line with most previous studies, Employment Outlook 1999 does not identify any significant effects of the measures of labour market rigidity on the average unemployment rate. There may be some effect on the employment rate, a result that chimes with the results reported in a recent paper by Di Tella and MacCulloch (1998). As noted by Nickell (1997) and by Nickell and Layard (1998), however, this could be a spurious correlation, since the southern European countries have the most rigid labour markets and lower female participation rates.

As regards dynamic effects, the study finds a negative correlation between rigidity and various measures of turnover.⁷ A positive correlation

⁷There is extensive debate on the comparison of turnover rates in the United States and Europe. In particular, contrary to what one might expect, gross annual job creation and destruction rates are similar in both areas (Boeri 1996). Bertola and Rogerson (1996) argue that this paradox is explained by the smaller wage dispersion in the European labour market, which would induce a greater degree of quantity adjustment in response to shocks. Boeri (1999) shows that this fact and the low flows in and out of unemployment that characterize rigid economies can be reconciled by considering the high incidence of temporary employment and large flows from job to job without going through unemployment in those economies. Comparing the US and Portuguese economies, Blanchard and Portugal (1998) show that while annual job creation and destruction rates are similar the United States has appreciably higher quarterly rates, indicating that firing restrictions inhibit temporary changes in employment but are less effective in the medium and long term.

is found between the average duration of a job and rigidity measures. Finally, the rates of entry to and exit from unemployment decrease as rigidity increases, and the duration of unemployment and the share of long-term unemployment both increase as the stringency of labour market legislation increases.

These conclusions fit the predictions of the theoretical model. The empirical analysis also offers another element that could not be obtained from the theoretical model, which assumes that labour is homogeneous. We showed that in a rigid system the unemployed tend to be unemployed for longer, while the employed are less likely to lose their jobs. Consequently, rigidities can have different effects on different segments of the labour force. The OECD study provides preliminary evidence for some such groups. Prime-age men, who are arguably the strongest segment of the labour force, have very similar employment (unemployment) rates across countries. These rates rise (decrease) with the degree of rigidity. By contrast, women and, above all, young people, i.e. the most vulnerable segments of the labour force, appear to be more exposed to the negative effects of labour market restrictions: youth unemployment, for example, is positively correlated with the level of rigidity. One possible explanation is that firing restrictions make employers more reluctant to hire first-time job seekers, since it is difficult to assess their “quality”. If they should prove to be unproductive, firms would have to bear the costs of firing them or keeping them employed. In other words, employers who hire marginal workers produce an information externality that contributes to their insertion in the production cycle. An inflexible labour market raises the cost of generating this information, slowing the entry of new workers into the economy and thus increasing youth unemployment. In the Italian case, however, this hypothesis runs up against the fact that there are especially flexible forms of contracts for new workers, such as apprenticeship contracts and, more recently, trainee contracts. The trend towards more flexible employment is spreading among the OECD countries, where in recent years there has been a sharp rise in fixed-term contracts and temporary jobs. One of the most important and stimulating lines of research in labour market studies is the effort to determine whether the persistence of high levels of youth unemployment, despite some slight improvement (OECD 1999), is due to market rigidities caused by inadequacies in the forms of contract that have been introduced, or instead has other roots.

In conclusion, in this class of models restrictions on dismissals do not generate appreciable direct effects on the unemployment rate, although they do

give rise to some effects on dynamic developments and flows. These predictions have empirical support, albeit with the many caveats that such analysis demands. In order to identify the effects of flexibility on employment, it therefore appears necessary to extend the scope of analysis along two lines: considering the indirect effects of flexibility in the labour market; and asking whether the concept of flexibility can be broadened beyond the simple idea of the degree of employment rigidity.

3 Firing restrictions and wage rigidity

In a recent survey for the Handbook of Labor Economics, Bertola (1998) highlights an empirical regularity that shows that countries with more stringent restrictions on dismissals tend to have high and undifferentiated wages and high unemployment rates. This observation suggests the importance of studying how employment protection legislation interacts with other institutional aspects of the labour market in the joint determination of wages and employment.

There is a high degree of complementarity between quantity restrictions and prices. One requires the other in order to be effective. If wages could be renegotiated in each period, in the event of a negative shock firing costs could be neutralised by offering a lower wage and thus inducing workers to abandon the market or bring labour costs back into line with labour productivity. The same holds in the opposite direction. Wage compression and firing restrictions provide strong protection for the wages of employed persons, even with high unemployment. Given a pool of unemployment and fixed wages for those with a job, if it were possible to fire workers in order to hire others the wage rigidity of the employed would be exposed to the pressure of competition from the unemployed. On the other hand, if it is illegal to offer a lower wage to a new worker for the same job, firms cannot use this channel to evade wage rigidity.⁸

These examples show that in order to be effective employment protection legislation must cover quantities (firing restrictions) and prices (rigidity and

⁸In addition to institutional factors, a further protection for the wages of the employed is more strictly economic in nature. As we have seen, in rigid labour markets unemployment lasts longer. As various scholars have argued (Blanchard and Summers 1986), the long-term unemployed gradually become less employable, hence the wage pressure they exert declines as the duration of employment increases. In other words, long-term unemployment does not trigger automatic reabsorption mechanisms, tending to persist instead (hysteresis).

wage compression). Bertola (1998) goes further, arguing that the combination of these two factors provides strong protection to the employed, increasing their bargaining power and therefore inducing higher wage growth in the presence of shocks than would otherwise occur. This view is supported by evidence on the similarity of employment rates among prime-age men across OECD countries, compared with sharp differences for other categories. The fact that young people and women tend to have lower employment rates in rigid labour markets suggests that labour market rigidities can have different effects on different groups. Verifying the accuracy of this hypothesis is an important line of future empirical and theoretical research. This work should focus on the joint determination of labour supply and demand, wage formation and the institutional characteristics of the labour market. The scope for empirical analysis is broad, as there is markedly little evidence available. In a study on labour and employment protection legislation, Lazear (1990) considers the possibility that the negative effects of employment protection measures are due to the potential correlation between labour protection and wage growth. The (weak) evidence he adduces against this possibility is based on the observation that, in a sample of 29 industrial countries over a period of 29 years, changes in employment protection legislation and minimum wage laws do not appear to occur together in any systematic manner.

One attempt at the empirical analysis of the interaction between the various aspects of the system in determining the performance of the labour market is offered by Blanchard (1998) and Blanchard and Wolfers (1999). Although it uses a reduced-form approach and therefore cannot be considered a true advance in the structural analysis of these issues along the lines proposed by Bertola (1998), this work underscores the importance of considering shocks and institutions together in explaining European unemployment. The authors argue that neither shocks nor institutional arrangements alone can explain the performance of the European labour market over the last three decades: shocks suffer from a cross-section difficulty, namely that they do not differ sufficiently among the countries to explain the different national situations, while institutions face a temporal problem, in that many of the institutions identified as potential causes of high unemployment already existed at the start of the 1970s, when the unemployment rate in Europe was very low. Moreover, the 1990s saw the emergence of signs of a reversal of the trend in employment protection legislation, without a corresponding halt in the rise in the unemployment rate in the leading European economies. Using panel regressions for a number of industrial countries, the authors show that the ability to explain the unemployment

rate increases sharply when the explanatory variables include interactions between macroeconomic shocks (such as total factor productivity and the interest rate) and measures of the institutional aspects of the labour market (such as the level of unemployment benefits, the degree of protection afforded those with jobs and the strength of labour unions). In short, it is necessary to examine how shocks are filtered by institutions in determining labour market performance. Their conclusions regarding future developments in European unemployment are moderately optimistic, since the negative shocks they identify are disappearing and institutions are slowly becoming less penalizing for employment. Regardless of the validity of the empirical results, there remain serious doubts about the soundness of these conclusions. First, a shock is by definition unpredictable, so it seems arbitrary to make claims about their future behaviour. Second, it is widely held that the industrial economies today are more dynamic and changeable than they once were. This conviction underpins the popularity of the concept of flexibility in the policy debate. All things considered, the same results can be used to arrive at a sharply different assessment - in the contemporary economic environment - of the adequacy of institutional arrangements that transmitted the effects of the negative shocks of past decades for such a long time.

4 Wage compression, geographical immobility and underemployment

The impossibility of adjusting wages to economic conditions carries a number of implications. For Italy, it is worth dwelling on the central determination of wage levels, with scant scope for local bargaining.⁹ This kind of wage rigidity can help explain the persistent high unemployment of the Italian South.¹⁰

Consider again the analytical framework set forth above. Instead of taking $A_i, i = \{l, h\}$ as different realizations of productivity in time, think of them as productivity levels in two regions, North and South, characterized

⁹The analysis concentrates on just one aspect of centralized bargaining, ignoring, for instance, the benefits of coordinated response to aggregate shocks. In a study of unemployment in OECD countries, Nickell (1997) finds a negative correlation between the level of coordination between unions and employers in wage-setting and the level of unemployment.

¹⁰See Faini (1997) for an analysis, in a context of growth, of the effects of centralized bargaining on convergence between regions at different stages of development in a model positing two different types of worker (skilled and unskilled).

by an immobile labour force (a point to which we shall return), where p represents the relative weight of the North in the economy. To illustrate the effect of wages on employment, one must explicitly introduce the labour supply. To keep the treatment at the simplest possible level, it is assumed that each individual chooses his level of consumption and of work so as to maximize:

$$\text{Max}_{\{c,l\}} U(c,l) = \log c - \frac{1}{2\alpha^2} l^2, \quad \text{s.t. } c = lw, \quad \alpha > 1 \quad (8)$$

Solving, we get $l^s(w) = \alpha$ and $c = \alpha w$. As we know, for the log function of utility in consumption, the income effect and the substitution effect perfectly offset one another, so the labour supply is independent of the wage rate. If wages are set locally, recalling that the first-order conditions of the firm imply $l_i^d(w) = A_i/w$, we get that employment is equal to α in both regions and that wages are equal to $\frac{A_i}{\alpha}$, so that the region with higher productivity has higher wages.¹¹

Now suppose that wages are the same in the two regions and are set as a weighted average of the wage levels that would be determined in each region separately:

$$\bar{w} = \frac{\lambda A_h + (1 - \lambda) A_l}{\alpha}, \quad \lambda \in (0, 1) \quad (9)$$

Compared with the previous situation, the wage level is lower in the North and higher in the South. Given that wages are pre-set, regional employment is determined by the lower of demand and supply at the given wage (Figure 1). Assume that if there is an excess supply of labour all workers are rationed in equal measure.¹² In the high-productivity region there is excess labour demand, so that employment is equal to α , while in the other there is excess supply, so that employment is determined by the labour demand at the prevailing wage:

$$L_l(\bar{w}) = L_l^d(\bar{w}) = \frac{\alpha A_l}{\lambda A_h + (1 - \lambda) A_l} < \alpha \quad (10)$$

The low-productivity region thus suffers from chronic, classic unemployment, because at the prevailing wage level there is an excess supply of

¹¹Using a utility function in which the income effect is stronger than the substitution effect, one gets that the high-productivity region also has both higher wages and a higher level of employment.

¹²This assumption appears realistic empirically, given the existence of social networks (whether family-based or public) that serve as insurance mechanisms between the employed and the unemployed.

labour. Unemployment is the higher, the greater the productivity differential ($A_h - A_l$) and the greater the weight of the high-productivity region λ in determining the wage rate \bar{w} . For example, if one assumes that the weights are equal to the relative sizes of the regional economies, so that $\lambda = p$, then unemployment is the greater, the greater the weight of the high-productivity region in the overall economy. The social cost in terms of the sum of the surplus of producers and consumers is equal to the area circumscribed by the points $\bar{L}_l^d L^s BD$ in Figure 1. Note, however, that if we take the model's implications literally, particularly as regards the assumptions on the production function, this cost is borne entirely by firms, in that labour demand elasticity of one implies that the wage bill does not change as w changes (in the graph, the two areas $\bar{L}_l^d L^s BA$ and $w_l AD\bar{w}$ are equal).

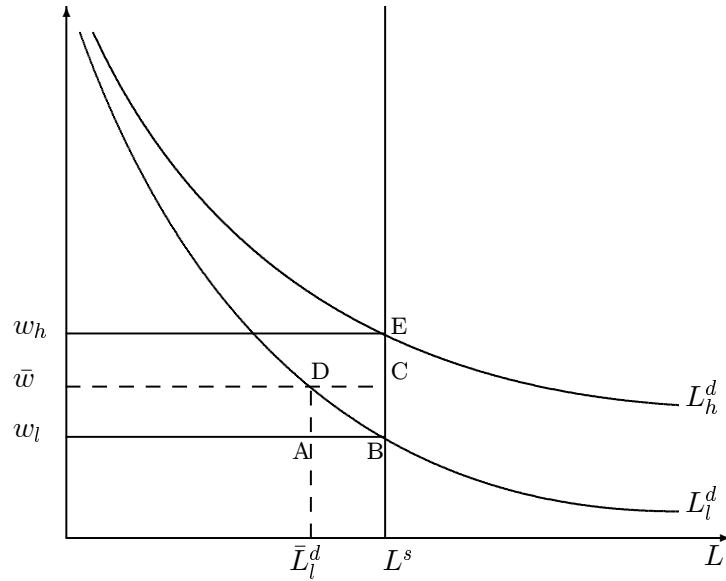
This mechanism has significant effects on the high-productivity region as well. Although the centralized setting of the wage rate implies that the total surplus (equal to the area under the curve L_h^d between zero and L^s) does not change, the lower wage rate results in a transfer of income from workers to firms equal to the area circumscribed by $\bar{w}CEw_h$. This observation implies that the firms in the high-productivity region might have an interest in fostering central wage determination.

To summarize, our example shows that central wage determination:

- generates unemployment in the South and reduces the total surplus in the economy;
- redistributes resources among groups, to the detriment of workers in the North and firms in the South.

The crucial assumption in this example is that labour is immobile. Given chronic unemployment in one region and full employment in the other, one would expect workers to move so as to restore market equilibrium, increasing the supply where there is excess demand and conversely. Geographical immobility is thus an essential of the high-unemployment equilibrium. This suggests that to explain the disappointing performance of the Italian labour market, we need to broaden the notion of rigidity from the narrow acceptations used to this point: specifically, we should examine the question of geographical mobility.

Figure 1: Employment and wages in segmented labor markets



5 Growth

This section deals with the relationship between rigidity and economic growth. Rigidity could influence the growth rate through a number of channels, in particular affecting innovative activity, the reallocation of productive factors, the adoption of new technology and investment. Surprisingly, this problem has been little studied in the framework of the enormous development of the new growth theory during the last decade. Empirically too, very little has been done — understandably, in part, given the difficulty of isolating the contribution of flexibility, itself hard to measure, to the growth rate of an economy. We here focus on one specific channel through which flexibility may affect growth, namely the investment process. The analysis is based essentially on just one work (Bertola, 1994), so the conclusions must be handled with great caution.

Section 2 showed that while curbs on firing do not have first order effects on the level of employment, they do tend to diminish profits for firms, which

cannot adapt optimally to changes in market conditions. This could have significant effects on the process of capital accumulation, hence on growth. To formalize this idea we must make some changes in the model used so far. First, the interpretation of parameter A needs to be further specified: rather than a simple productivity shock, we now read it as the level of an accumulable factor, in the tradition of AK endogenous growth models. This factor can be usefully referred to as capital.¹³ In every period, output depends both on the level of A accumulation and on the realization of a technological shock: $A_i = A * i$, $i = \{h, l\}$, $h > l$, where $\Pr\{i = h\} = p$. Second, we assume that output is generated not via an aggregate production function but a continuum, of mass one, of production sites, all with the same endowment of the accumulable factor A .¹⁴ Shocks are distributed independently both of time and of plants. Aggregate output is:

$$Y = \int_0^1 A * i_j * \log L_j dj \quad (11)$$

where L_j is the quantity of labour employed at the sites that have been affected by technological shock j . As there is a continuum of sites, there is no uncertainty at aggregate level. In every period, the share of high-productivity sites is equal to the probability that the shock will take the value h , so that aggregate output can be rewritten as:

$$Y = A * (ph \log L_h + (1 - p)l \log L_l) \quad (12)$$

Workers are formalized as in the previous section. On this basis, we assume that wages cannot differ from plant to plant and that they depend on average productivity:

$$w(A) = \frac{pA_h + (1 - p)A_l}{\alpha} = \frac{A(ph + (1 - p)l)}{\alpha} \quad (13)$$

¹³The literature identifies the factor with constant returns to scale as technical progress. The results of this section should accordingly be considered to be applicable more generally than just to the concept of physical capital but also to such cumulable factors as technological innovation and human capital.

¹⁴The homogeneous allocation of capital among plants could be derived endogenously assuming that the allocation is decided before the technological shock takes place, in accordance with the relative irreversibility of investment and the lag characterizing the investment process, and that returns to scale are decreasing for the individual plant but constant at aggregate level. This generalization would in no way alter the results but would bring a further degree of complexity, so it is not utilized here.

Finally, we assume that there exists a “representative entrepreneur” who owns the accumulable factor and decides the accumulation path. The absence of aggregate uncertainty means that he solves a deterministic problem in terms of saving, which greatly simplifies the analysis.¹⁵ The entrepreneur does not offer labour and has logarithmic consumption preferences:

$$\text{Max}_{\{c_t\}_{t=0}^{\infty}} E\left\{\sum_{t=0}^{\infty} \beta^t \log c(t)\right\} \quad (14)$$

$$\text{s.t.} \quad A(t+1) = (1+r)A(t) - c(t) \quad (15)$$

where, given that there is no aggregate uncertainty, the return to capital rA (which coincides with average profit) is a deterministic variable, even where there are shocks at the level of the single production plant. The solution to this problem is well known: the Euler equation that governs the dynamics of consumption is of the form:

$$c(t+1) = \beta(1+r)c(t) \quad (16)$$

For a given initial value $A(0)$, the accumulable factor and consumption follow the paths:

$$A(t) = [(1+\bar{r})\beta]^t \frac{1}{(1-\beta)^2} A(0) \quad (17)$$

$$c(t) = [(1+\bar{r})\beta]^t \frac{1+r}{1-\beta} A(0) = (1+\bar{r})(1-\beta)A(t) \quad (18)$$

As long as $(1+\bar{r})$ is larger than the inverse of the intertemporal discount rate $1/\beta$, there will be an indefinite growth in consumption and in the stock of capital, the classical result of endogenous growth models. Given that the rate of growth of the accumulable factor is $(1+\bar{r})\beta - 1$, it follows that

$$\frac{\partial(\dot{A}/A)}{\partial\bar{r}} = \beta > 0 \quad (19)$$

The rate of capital accumulation is thus increasing in the expected return to capital.

¹⁵Equivalently, one may assume a multiplicity of entrepreneurs, each with the same share of the accumulable factor. In this case, one would need the additional assumption that capital markets are complete and thus that entrepreneurs can completely diversify their idiosyncratic risk, for instance by holding shares in a portfolio composed of all the sites within the economy.

Now let us consider flexible and rigid economies (f, r) . The appendix shows that in both types, profits are a linear function of A , so that the average and the marginal return are independent of the level of A . So if we denote returns to capital with \bar{r}_f in the flexible and \bar{r}_r in the rigid economy, the fact that expected profits are lower in the latter, as is demonstrated in Proposition 1 in the appendix, implies that $\bar{r}_f > \bar{r}_r$.¹⁶ Using this result and equation 19, we reach an important conclusion: the rigid economy grows more slowly than the flexible economy. In this simple model, the rigidity of the labour market limits entrepreneurs' capacity to make profits, adversely affecting the rate of growth of the accumulable factor, hence that of the economy as a whole. Along a balanced growth path, in fact, the rate of growth in output and in wages is equal to that of the accumulable factor: $\dot{w}/w = \dot{y}/y = \dot{A}/A$.

Bertola (1994) obtains the same result in a more general model. Here, firing restraints have an impact effect that implies greater labour intensity, hence a redistribution of income in favour of workers. The conclusions change, however, when the general equilibrium effects on the process of accumulation are considered. The cost of dismissals reduces the growth rate of the economy and of wages, adversely affecting the welfare of all agents.

Caballero and Hammour (1998) reach similar conclusions from different premises. This focus on the concept of "appropriability," i.e. the fact that given specific investments, tied to the characteristics of the other factor and not immediately redeployable elsewhere, one of the two parties (workers or employers) can appropriate a larger share of the output than was originally "agreed" when the investment decision was made. In the authors' opinion, the labour legislation passed in the late '60s and early '70s augmented workers' bargaining power, increasing their appropriation capacity and generating a temporary rise in labour's share of output. In the short run, in fact, the supply of capital is inelastic and the elasticity of substitution between the two factors is low, so there is little scope for reducing this appropriation. In the long run, however, substitution elasticity is considerably greater, as firms can choose among various production technologies. They consequently tended towards less labour-intensive technology, substituting capital for labour and reducing the latter's share of total output.¹⁷ Such a

¹⁶As we have seen, this result depends on the fact that in the flexible economy labour is allocated more efficiently, assigning a higher share of it to the high-productivity sites, in order to equalize the marginal productivity of all plants and thus maximize output.

¹⁷These observations are consistent with the trends in capital formation and employment in Europe and the United States described in Caselli et al. (1999).

process may have repercussions on growth, because the technology used may not be the most efficient, and once the dynamic has played through it can result in a reduction of labour's share to below the original level. The authors contend that this is what actually happened with the wage moderation marking the European economies in the '90s.

Caballero and Hammour thus emphasize the role played by the rules governing the division of the surplus. In the long term, rules that favour the appropriation of rents by one factor will induce a shift in the other factor so as to reduce that possibility. Their model assumes that a specific (hence appropriable) investment is made by the firm, and that rigidity favours appropriation by workers. But this argument can be reversed. One of the most commonly cited justifications for firing restraints (OECD, 1994) is that they make for stable relations between the worker and the firm, thus giving the worker an incentive to invest in specific human capital because he is protected against undue appropriation by the employer. Moreover, given imperfect capital markets, firing restraints attenuate fluctuations in individual income and thus increase workers' welfare by stabilizing consumption over time. Given imperfect capital markets, such firing restraints as severance pay could well form part of an optimal employment contract. Restraints on dismissals represent a form of insurance for risk-averse workers provided by a risk-neutral firm, in the spirit of "implicit contracts" (Rosen, 1985). This model of industrial relations is said to underlie systems characterized by stable employment, such as the German and Japanese economies, in contrast to the United States. The latter's better performance in the '90s should not cause us to forget the different assessment of the relative merits of the two types of industrial organization in the previous decade.

5.1 Transition and steady-state

One of the main problems in explaining high European unemployment as the result of poor labour market flexibility is temporal in nature. As Blanchard (1998), among others, has observed, the European unemployment rate was relatively low during the '70s, when the bulk of the present restrictions were already in place, and peaked in the '90s when measures that increased flexibility were introduced.

This section offers a possible explanation of the contradiction.¹⁸ The introduction of the firing curbs in the early '70s unquestionably helped blunt

¹⁸For an overview of the explanatory hypotheses see Sestito (1997).

the effect of the oil shocks on jobs, producing less unemployment than would have been the case otherwise. But the general, long-term effects entailed a gradual rise in the equilibrium rate of unemployment. The firing restrictions themselves were one reason that this new equilibrium was reached only gradually; they slowed the decrease in employment and prolonged the process to later years.

As for the rise in unemployment in the '90s, Bertola and Ichino (1995) suggest an interesting thesis. In their view, the introduction of more flexibility in Europe in the last decade was halting and uncertain. Consequently, while firms whose need was to downsize apparently took advantage of the new opportunities (as attested, for instance, by the sharp rise in job destruction in Italy in 1993), those interested in expanding staff may have elected not to do so because of uncertainty over the credibility of the process: if flexibilization were to be halted or reversed once the new workers were hired, they would have trouble coping with negative shocks. In this view, that is to say, the uncertain reform process left the labour market in “midstream,” greater flexibility having triggered the destruction of redundant jobs but not the creation of new ones.

Job creation takes time and resources, unlike job destruction, which in the absence of institutional constraints can be instantaneous (Mortensen and Pissarides, 1994). The transition to a more flexible system, though potentially benefiting steady-state growth, can have adverse effects in the present, as overstaffed firms will dismiss workers immediately whereas expanding ones will hire only gradually. Bertola and Ichino (1995) note that this negative effect may be longer-lasting, the less credible the process of labour market reform.

6 Conclusions

This paper sums up the state of the theoretical debate on the role of labour market flexibility in determining the performance of the European economies. It has been shown that the “direct” effects are important in determining cyclical variations in employment, in the duration of unemployment and in the turnover rate but do not appear capable of accounting for high European unemployment. The analysis has accordingly been extended in two directions: considering the indirect effects of rigidity and generalizing the concept of flexibility in an economy.

The following conclusions on indirect effects have been reached:

A. One possible explanation of high unemployment is that rigidity in the labour market tends to be accompanied by fast growth in real wages, in that restraints on firing and the attached rigidities increase the bargaining power of employed workers, protecting them from the downward wage pressure exerted by the jobless.

B. The longer duration of unemployment may have important effects, both because the long-term unemployed tend not to exert the downward pressure on wages that could contribute to their employment (hysteresis) and because it may act as a significant entry barrier vis-à-vis young workers, a major problem in Europe.

C. Given geographical immobility, the setting of uniform wage levels depresses employment in the low-productivity region and handicaps firms in that region and workers in the high-productivity one.

D. Rigidity may adversely affect the process of capital accumulation and growth, in connection with lower returns to capital.

A particularly important theme for future research is the extension of the concept of flexibility to other aspects of the economy, not restricting it to the labour market alone. A list of topics for further study should include the flexibility of human capital, which depends on the level and quality of the education provided by the school system; labour supply, bound up with geographical mobility, hence with the efficiency of the housing market and with social protection and family networks; the process of market entry and exit on the part of businesses and self-employed workers, which depends among other factors on the efficiency of government, infrastructural endowments and professional constraints that act as entry barriers; the system of corporate governance; and the credit market, especially the financing of innovative activity.

Practical strategies for the reform of the economy in general and the labour market in particular must then be formulated. There is a substantial body of literature on these issues, from which we select two key indications¹⁹:

- Because it is potentially harmful to the interests of specific groups with powerful political and social representation, the reform of the labour market faces tenacious resistance. However, there are measures that can alleviate the problem, providing greater representation for the groups that are most damaged by the status quo. For example, fixed-term employment contracts may help enact a “two-stage reform” (Saint-Paul, 1998), in which gradually,

¹⁹See among others Saint-Paul (1998) for an analysis of the political economy of labour market reform and Snower and Dehesa, eds. (1997) for effective reform measures.

as the unemployed are hired on fixed-term contracts, their power increases.

- There are significant complementarities in the reform process. According to Orzag and Snower (1998), labour market reforms are characterized by great complementarity economically (they are more effective if introduced simultaneously) and politically (easier to pass if introduced all at once). For instance: supply-side measures will be more effective when accompanied by policies to stimulate demand; programmes for the hiring of the long-term unemployed will produce better results if unemployment benefits are designed so as to maximize the incentives for the jobless to return to the market; the approval of measures to enhance labour market flexibility will encounter less resistance when accompanied by the introduction of an appropriate system of unemployment benefits, which is still lacking in Italy.

These considerations are especially important. The point is that effective reform of the economy will require tackling a whole series of interrelated matters simultaneously. This is a worthy theme for further study.

7 Appendix: Proofs

PROPOSITION 1: Output and profits are lower in the rigid economy.

Proof

The demonstration that average profits are higher in the flexible economy is immediate. By definition, L^f is the level of labour that corresponds to the sole maximum of the profit function in each state:

$$L^f(i) \equiv \text{Argmax}_L \{A_i \log L - wL\} \quad (20)$$

Since $L^r \neq L^f(i)$, $i = \{h, l\}$, profits in each state are higher in the flexible economy, and so also are average profits.

In terms of output, define $A_h \equiv \theta A_l$, $\theta > 1$. We want to show that $\forall p \in (0, 1)$

$$p\theta A_l \log\left(\frac{\theta A_l}{w}\right) + (1-p)A_l \log\left(\frac{A_l}{w}\right) > (p\theta A_l + (1-p)A_l) \log\left(\frac{p\theta A_l + (1-p)A_l}{w}\right) \quad (21)$$

Rearranging, this expression reduces to:

$$p\theta \log \theta > (p\theta + (1-p)) \log(p\theta + 1-p) \quad (22)$$

Consider two functions $b(p).c(p)$, defined respectively by the right-hand and left-hand sides of the previous inequality. It can be immediately verified that:

1. $b(0) = c(0); \quad b(1) = c(1)$.
2. $b'(p) = \theta \log \theta > 0$.
3. $c'(p) = (\theta - 1) \log(p(\theta - 1) + 1) + \theta - 1 > 0$.
4. $c''(p) = \frac{(\theta - 1)^2}{p(\theta - 1) + 1} > 0$.

In other words, $b(p)$ is linear and increasing, while $c(p)$ is increasing and strictly convex. Since the two have the same value for $p = 0, p = 1$, from the convexity of $c(p)$ it follows that $b(p) > c(p) \quad \forall p \in (0, 1)$.

PROPOSITION 2. On the assumptions of Section 4, profits are linear in A for both economies.

Proof

Recall that wages are equal to $w(A) = \frac{A(ph + (1-p)l)}{\alpha}$. Consider the flexible economy. The quantity of the labour factor utilized at a productive site i is equal to:

$$L(A, i, w) = \frac{Ai}{w} = \frac{\alpha i}{ph + (1 - p)l} \quad (23)$$

Substituting 23 into the profit function, we get

$$Ai \log L(A, i, w) - wL(A, i, w) = Ai \log\left(\frac{\alpha i}{ph + (1 - p)l}\right) - Ai \quad (24)$$

For the rigid economy, we have

$$L(A, w) = \frac{A(ph + (1 - p)l)}{w} = \alpha \quad (25)$$

Substituting this into the profit function, we get:

$$Ai \log L(A, w) - wL(A, w) = Ai \log \alpha - A(ph + (1 - p)l) \quad (26)$$

In each case, total profit is a linear combination of the profits at the high-productivity and low-productivity sites.

References

- [1] Alvarez F. - Veracierto M. (1998), Search, Self-Insurance and Job-Security Provisions, mimeo, Federal Reserve Bank of Chicago.
- [2] Bertola G. (1999), Microeconomic Perspectives on Aggregate Labor Markets, in *Handbook of Labor Economics vol.3B*, ed. da O.Ashenfelter - D.Card, Amsterdam, North-Holland, pp.2985-3028.
- [3] Bertola G. (1994), Flexibility, Investment and Growth, *Journal of Monetary Economics*, vol. 34, pp. 215-238.
- [4] Bertola G. - Ichino A. (1995), Crossing the River: a Comparative Perspective on Italian Employment Dynamics, *Economic Policy*, n. 21, pp. 359-420.
- [5] Bertola G. - Rogerson R. (1996), Institutions and Labor Reallocation, *Temi di discussione della Banca d'Italia*, n. 272.
- [6] Bentolila S. - Bertola G. (1990), Firing Costs and Labor Demand: How Bad is Euroclerosis?, *Review of Economic Studies*, vol. 57, pp. 381-402.
- [7] Blanchard O. (1998), European Unemployment: Shocks and Institutions, Baffi Lecture tenuta alla Banca d'Italia.
- [8] Blanchard O. - Wolfers J. (1999), The Role of Shocks and Institutions in the Rise of European Unemployment: the Aggregate Evidence, mimeo, MIT.
- [9] Blanchard O. - Portugal P. (1998), What Hides behind the Unemployment Rate: Portugal and US, NBER Working Paper, n. 6636.
- [10] Blanchard O. - Summer L. (1986), Hysteresis and the European Unemployment Problem, *NBER Macroeconomics Annuals 1*, pp. 15-78.
- [11] Boeri T. (1999), Enforcement of Employment Security Regulations, on-the-job Search and Unemployment Duration, *European Economic Review*, vol. 43, pp. 65-89.
- [12] Boeri T. (1996), Is Job Turnover Countercyclical?, *Journal of Labor Economics*, vol. 14, pp. 603-25.

- [13] Caballero R. - Hammour M. (1998), Jobless Growth: Appropriability, Factor Substitution and Unemployment, *Carnegie-Rochester Series on Public Policy*, vol. 48, pp. 51-94.
- [14] Caselli P. - Pagano P. - Schivardi F. (1999), Investment and growth in Europe and the United States in the nineties, mimeo, Banca d'Italia.
- [15] Di Tella R. - MacCulloch B. (1998), The Consequences of Labor Market Flexibility: Panel Evidence Based on Survey Data, mimeo, Harvard Business School.
- [16] Faini R. (1997), Trade Unions and Regional Development, Centro Studi D'Agliano, Working Paper, n. 115.
- [17] Garibaldi P. (1998), Job Flow Dynamics and Firing Restrictions, *European Economic Review*, vol. 42, pp. 245-275.
- [18] Hansen G. (1985), Indivisible labor and the Business Cycle, *Journal of Monetary Economics*, vol. 16, pp. 309-327.
- [19] Hopenhayn H. - Rogerson R. (1993), Job Turnover and Policy Evaluation: a General Equilibrium Analysis, *Journal of Political Economy*, vol. 101, pp. 915-938.
- [20] Lazear E. (1990), Job security provision and employment, *Quarterly Journal of Economics*, vol. 105, pp. 699-726.
- [21] Lijunqvist L. (1997), How do Layoffs Costs Affect Employment, Working Paper, Stockholm School of Economics.
- [22] Mortensen D. - Pissarides C. (1994), Job Creation and Destruction in the Theory of Unemployment, *Review of Economic Studies*, vol. 61, pp. 379-415.
- [23] Nickell S. (1997), Unemployment and Labor Market Rigidities: Europe vs North America, *Journal of Economic Perspectives*, vol. 11, pp. 55-74.
- [24] Nickell S. - Layard R. (1998), Labor Market Institutions and Economic Performance, versione preliminare per l'*Handbook of Labor Economics*, vol. 3.
- [25] OCSE (1999), Employment Outlook 1999, OCSE, Parigi.

- [26] OCSE (1994), *The OECD Job Study: Facts, Analysis, Strategies*, OCSE, Parigi.
- [27] Orszag J. - Snower D. (1998), *Anatomy of Policy Complementarities*, CEPR Discussion Paper, n. 1963.
- [28] Pissarides C. (1985), *Short Run Equilibrium Dynamics of Unemployment, Vacancies and Real Wages*, *American Economic Review*, vol. 75, pp. 676-690.
- [29] Rogerson R. (1988), *Indivisible Labor, Lotteries and Equilibrium*, *Journal of Monetary Economics*, vol. 21, pp. 3-16.
- [30] Rosen S. (1985), *Implicit Contracts: a Survey*, *Journal of Economic Literature*, vol. 23, pp.1144-75.
- [31] Saint-Paul G. (1998), *The Political Economy of Firing Costs*, mimeo, Universitat Pompeu Fabra.
- [32] Sargent T. (1987), *Dynamic Macroeconomic Theory*, Harvard University Press, Cambridge, Ma.
- [33] Sestito P. (1997), *La questione del lavoro e le interpretazioni proposte*, in *Disoccupazione di fine secolo*, ed. da P. Ciocca, Bollati Boringhieri.