# The economics of unfair dismissal

# in the United Kingdom,

# and other topics in public policy

Ioana Elena MARINESCU

London School of Economics and Political Science

Thesis presented for the award of a PhD in economics in 2006

Supervisors: Professors Stephen Nickell and Alan Manning

## Declaration of authenticity

I, Ioana Elena Marinescu, certify that I am the author of the work presented in this thesis, with the exception of chapter 4, which was co-written with Philippe Aghion. I have contributed 50% to chapter 4.

I also declare that the paper "Are Judges Sensitive to Economic Conditions? Evidence from UK Employment Tribunals", which forms chapter 3 of this thesis, was in some slightly different form included in my French PhD (EHESS, Paris 2005). Some other small parts of the thesis were also included in that work, but in a very preliminary state of completion, so that the thesis currently submitted can be considered, for more than 70% of it, to be new work.

Ioana Marinescu.

I. Marinescu

# Thesis abstract

Workers and firms face substantial uncertainties about their prospects in the labor and product markets. The first three chapters of this thesis analyze how firing costs affect firms' behavior and workers' outcomes in the face of uncertainty about match quality and changing economic conditions. In the final chapter, I show how macroeconomic policy can reduce the risks associated with changing economic conditions.

First, I examine a 1999 UK reform that lowered from two years to one year the tenure necessary for a worker to be able to sue their employer for unfair dismissal. After the reform, we observe a significant decrease in the firing hazard for workers with zero to two years tenure relative to the control group, and no overall increase in unemployment. Using a simple model based on the assumption that firms learn about match quality over time, I show that the empirical results are consistent with increased match quality after the reform.

Second, I generalize the simple model developed in the first chapter. In particular, I allow for match quality to change over time. The model is useful to understand and predict how firing costs and various forms of uncertainty affect the separation hazard.

Thirdly, I analyze the implementation of unfair dismissal legislation by judges in the UK. Judges seem to compromise between workers' and firms' interests. If workers are unemployed, judges decide more often in their favour when unemployment rates are higher. The reverse is true when workers have found a new job.

Finally, in work co-authored with Philippe Aghion, we examine whether the government borrowing and spending more in recessions can increase growth by relaxing economic agents' credit constraints. Using a panel data of OECD countries, we find that indeed countercyclical public debt policy is more growth enhancing when private credit is less abundant.

# Table of contents

# Acknowledgements

I would like to thank my supervisors Stephen Nickell and Alan Manning for their support. In particular, I would like to acknowledge the contribution of Stephen Nickell, who has been providing me with crucial intellectual and moral support; his criticisms and encouragement were invaluable. Finally, I would like to thank Philippe Aghion, whose double role as a guide and a collaborator has been essential to my professional development.

# Introduction

Two of the most important uncertainties faced by economic agents are firms' uncertainties about the demand for their products and workers' uncertainty about the demand for their labor. When negative shocks occur, firms risk going bankrupt and workers becoming unemployed. As long as agents have sufficient cash or the ability to borrow to cushion negative shocks, this is not a major concern. Often enough, however, considerations of credit constraints and limited ability to pay play an important role, making the provision of insurance through private or public means desirable.

It has been long recognized that workers are typically more financially constrained than firms. This makes it potentially efficient to make firms participate in the insurance of workers against job loss. One of the ways of providing insurance against job loss is to mandate firms to compensate workers who are dismissed for no fault of their own. If workers sue their employer for unfair dismissal and the ground for the dismissal is found to be "unfair", then employers must pay an even greater compensation. The imposition of such firing costs may however conflict with the objective of insuring firms against bad shocks, since job destruction is more likely to occur when firms face adverse economic conditions. Wouldn't the imposition of firing costs on firms who already face adverse economic conditions lead to more bankruptcies and ultimately more workers losing their jobs?

Firing costs are however but one aspect of the costs faced by firms when economic conditions are deteriorated. Another important difficulty is that credit constraints are usually not constant over the cycle; instead, firms typically find it harder to borrow precisely when the economy is doing badly. This implies that, during a recession, firms are limited in their ability to invest in order to develop new products and technologies, and catch up on profits once the economy recovers. The government can improve this situation by borrowing and spending more when economic conditions are bad. This can release firms' credit constraints through various channels, such as demand stimulation or investors' confidence building.

This thesis addresses some of the questions raised by firing costs and countercyclical macro policies. In the first three chapters, I analyze the economic impact of firing costs, focusing on firms' behavior, workers' outcomes, and judges' decisions about whether to impose firing costs on firms. In the last chapter, in work co-authored with Philippe Aghion, I analyze how credit constraints affect the impact of countercyclical macro policies on economic growth.

From a theoretical perspective, firing costs have an ambiguous effect on productivity and employment. On the employment side, firing costs could either increase or decrease employment. Indeed, firing costs dissuade firing and thus partially insure workers against the prospect of being fired. At the same time, firing costs can reduce hiring: if a firm knows that in the future it may have to fire a worker and pay firing costs, it is more reluctant to take on that worker today. Firing costs also have an ambiguous effect on productivity. On the one hand, in as much as they provide insurance, they can motivate the worker to invest in specific human capital. On the other hand, because firing costs make the threat of dismissal less credible, workers may work less hard and/or demand higher wages, which would have a detrimental effect on productivity and/or employment. Adding to these theoretical ambiguities, the empirical literature has not been able to come to a strong conclusion regarding the effects of firing costs on either employment or productivity.

This thesis contributes to this literature both from an empirical and a theoretical perspective, filling two important gaps in the previous literature. First, the literature has been neglecting the fact that firing restrictions are typically only imposed for workers that have reached a given tenure with their current employer, i.e. there is a probationary period during which employers can fire at will[1]. Such is the case in particular in the United Kingdom, where unfair dismissal legislation requires workers to have been continuously employed for at least one year in order to be entitled to claim unfair dismissal. In chapters 1 and 2, I analyze, both theoretically and empirically, how the introduction of a probationary period affects firms' hiring and firing behavior. I also investigate empirically the impact of such a probationary period on workers' outcomes in the labor market (chapter 1). The theoretical analysis assumes that firms learn about match quality by observing the worker's behavior over time. Moreover, match quality is allowed to evolve over time (chapter 2). Compared to no firing costs, the introduction of a probationary period[2] increases the firing hazard (i.e. the probability that a worker is fired at some tenure t given that they have not been fired before) just before the end of

---

[1] Depending on the specific country and legislation, there may exist some restrictions to firing during the probationary period, but these restrictions are typically much less stringent than after the end of the probationary period.
[2] I.e. no firing costs till the end of the probationary period and positive firing costs thereafter.

the probationary period, and decreases it afterwards. Chapter 2 also investigates the effects of various forms of uncertainty about match quality on firms' optimal firing strategy, and on the firing hazard. The empirical analysis in chapter 1 draws on a change in British law in 1999, whereby the probationary period was reduced from two years to the current one year. After the reform, we observe a significant decrease in the firing hazard for workers with zero to two years tenure relative to workers with two to four years tenure, and no overall increase in unemployment. The calibration of the theoretical model reveals that firms recruit better workers after the reform, and also monitor workers somewhat better. Hence the results are consistent with an increase in match quality after the reform.

Second, the literature on firing costs has typically been concentrating on *de jure* legislative provisions, largely overlooking the fact that the concrete implementation of the law by judges may matter just as much as the letter of the law. Thus, in chapter 3, I look at the determinants of judges' decisions in unfair dismissal cases in the UK, focusing in particular on the influence of economic conditions. The effect of economic conditions on the implementation of firing costs is important in as much as it can lead to firing costs being de facto pro or counter cyclical, even though they may have not been explicitly designed to depend on economic conditions. I find that a higher unemployment or bankruptcy rate makes judges more likely to decide in favor of firms, thus exempting them from firing costs. Only if the worker is unemployed are judges more likely to decide in the worker's favor when unemployment rates are higher. These findings are consistent with firing costs being made more pro-cyclical by judges' decisions, and with judges weighing firms' and workers' welfare when making their decision. In particular, judges seem to be sensitive to the fact that firms are credit constrained and thus have a limited ability to pay during recessions.

While judges taking into account firms' credit constraints may contribute to limiting the risk of bankruptcy, this effect cannot be very large. Indeed, judges' actions cannot do much about the existence of a recession or credit constraints in the first place. To limit the negative effect of recessions, macroeconomic policy has long been seen as the tool of choice. In particular, there has been a long established tradition of Keynesian inspiration according to which countercyclical macro policy can enhance growth: the idea is that the state should stimulate demand during recessions to alleviate the adverse effects of these recessions and encourage recovery. More recently, there has been a growing skepticism as to whether such policies are really efficient. In particular, it has been argued that macroeconomic policies can at best have a short run effect but that long term growth is governed by institutions. However, this debate has been neglecting two important issues. First, such claims of efficiency or inefficiency of countercyclical macro policies need to be not only backed by theory but also tested empirically. The ability to perform meaningful empirical tests has however been limited by the challenges posed by the measurement of the countercyclicality of macro policies. Second, the efficiency of countercyclical macro policies may depend on a number of factors, and in particular the degree of financial constraints faced by the economy, i.e. the degree of financial development. Indeed, if credit constraints are a serious problem and firms are more credit constrained during recessions, then investments in new products and technology are limited during recessions, precisely when their opportunity cost is lower. Thus, a countercyclical macro policy, by supporting firms during recessions, can encourage investments that enhance long-run economic growth. My paper written with Philippe Aghion (chapter 4) sheds new light on these questions. We first address the challenging issue of providing yearly measures of the countercyclicality of debt policy. We use a series of different measures and provide a new methodology for estimating the cyclicality of debt policy. Second, we analyze whether the effect of countercyclical debt policy depends on financial development and find that this is indeed the case. Specifically, the less private credit an economy is able to count upon, the more growth-enhancing is countercyclical public debt policy. This has important consequences for policy making in the European Monetary Union: thus, our results imply that making public debt policy more countercyclical could increase economic growth in the EMU.

The thesis is structured as follows. In chapter 1, I focus on the analysis of the impact of a British reform that lowered from two years to one year the tenure necessary for a worker to be able to sue their employer for unfair dismissal. In chapter 2, I analyze a general model of relationship dissolution, focusing in particular on the impact of uncertainty and separation costs on the hazard of separation. Chapter 3 analyzes the determinants of judges' decisions in unfair dismissal cases in the United Kingdom. Finally, chapter 4 explores the role of financial development in explaining the impact of countercyclical macro policies on economic growth.

# Chapter 1

# Shortening the Tenure Clock: The Impact of Strengthened U.K. Job Security Legislation[1]

Ioana Marinescu

Abstract:

Even in countries with stringent job protection, workers typically only benefit from job security once they have worked at their employer beyond a minimum qualifying (or probationary) period. This paper analyzes how such a probationary period influences firms' behavior and workers' outcomes. I specifically examine a 1999 British reform that lowered from two years to one year the tenure necessary for a worker to be able to sue their employer for unfair dismissal. I first construct a model based on the assumption that firms learn about match quality over time. The model predicts that, after the reform, the hazard of firing workers between 1 and 2 years tenure decreases relative to the hazard beyond 2 years in all cases. Moreover, if, to avoid keeping lemons beyond the shorter qualifying period, firms react by recruiting workers more carefully, the hazard between 0 and a few months is predicted to decrease relative to the hazard beyond 2 years; an increase in monitoring has the opposite effect. Cox proportional hazard regressions show that the reform decreased the firing hazard between 0 and 2 years relative to the hazard between 2 and 4 years by about 30%. The calibration of the model reveals an increase in both recruitment and monitoring efforts, hence match quality. Consistent with an increase in match quality, I find that low tenure workers are more likely to receive training after the reform. Lastly, the reform has no detectable impact on unemployment duration, wages or employment.

---

[1] Parts of this paper at a much earlier stage were included in my PhD defended in June 2005 at the EHESS, Paris.

# 1 Introduction

US "employment at will" – the right for employers to dismiss workers whenever they want and for whichever reason, i.e. "at will" – is often contrasted with European job security legislation. In particular, job security is commonly portrayed as one of the causes of high unemployment and slow growth in Europe. However, the difference between US and European job security legislation is not quite as stark as it would seem at first glance. For example, in the US, there are quite a few exceptions to the employment at will rule. Some of them are due to law and jurisprudence, such as anti-discrimination laws, and others to custom, such as the institution of tenure in US universities. Still, the majority of the workforce in the US remains under "employment at will". By contrast, in Europe, and in most developing countries, employers can generally only fire workers for a "fair" reason. However, it is usually not the case that workers benefit from such job security from day one of the employment relationship. Instead, they are only granted full job protection rights once they have worked for their employer for the full length of a probationary period. Even in countries with high firing costs, dismissal costs are thus usually very low in the beginning of the employment relationship, and they significantly increase with tenure.

Conditioning employment protection on workers having reached a given tenure can be seen as a way to tackle the trade-offs generated by firing costs, combining the best of employment at will and job security. Indeed, on the one hand, firing costs may reduce the burden of economic downturns by making firms internalize the social costs of firing. Moreover, firing costs can increase productivity either by resulting in better job matching or by stimulating investment in human capital (Malcomson 1999). And, for risk averse workers, job security is a benefit in itself. On the other hand, higher firing costs will tend to reduce hiring in as much as they increase the cost of labor (for a theoretical illustration of the trade-off, see Bertola (1992)). High firing costs may also prevent the sorting of workers into the jobs they are best suited to, thus reducing productivity (Blanchard and Katz 1997).

A probationary period mitigates the latter problem, since firms can fire workers unsuited to the job at low cost at the beginning of the employment relationship. The institution of a probationary period is also related to the "last in, first out" rule, which requires that, when a firm lays off workers, it should first lay off those with lowest tenure on the job. This rule allows firms to adjust their workforce at lower cost, while preserving most workers' job security. Tenure-dependant job protection is thus a measure that can balance workers' and firms' objectives.

This paper analyzes a specific example of a probationary period provision in the United Kingdom. The right for dismissed workers to sue their employer for unfair dismissal is only granted after a given tenure on the job: before June 1999, this required tenure was two years, and after June 1999 it was reduced to one year. This source of variation allows me to shed light on two questions. First, what are the effects of having such a probationary period[2] on firms' firing behavior? Second, what is the impact of a reduction in the probationary period on firms' personnel management practices and workers' labor market outcomes? The answers to these questions are of particular interest in the context of European employment policies. Indeed, many European countries developed fixed-term contracts to allow for a probationary period without directly altering their protective legislation, and France, taking a further step, introduced in August 2005 a new employment contract, the CNE ("contrat nouvelles embauches", i.e. "contract for new hires"). The latter allows firms with less than 20 employees to benefit from a 2 years probationary period during which employment is almost at will, while standard job protection is granted after the end of the probationary period. In the early days of 2006, the French government proposed to extend the CNE, allowing all firms to hire employees below 26 years old under a CNE type contract. This was named the CPE ("Contrat Premier Emploi", i.e. first job contract). The CPE was seen by lots of people as a step towards complete liberalization of the labor market and was thus opposed by millions of demonstrators. As a result, the proposal was not surprisingly withdrawn. The

[2] From a strictly legal point of view, the change in unfair dismissal rights is not equivalent to a change in what is legally defined as the probationary period (which in fact plays a very minor role in UK law). But this terminology is useful to conceptualize the problem.

German government led by Angela Merkel also plans to increase the probationary period from 6 months to 2 years, but the law has not yet been enacted.

A large and well-established body of literature relates firing costs and employment across countries (Djankov et al. 2004) or across countries and time (Lazear,1990, OECD 1999, Heckman and Pagès 2003, Nickell, Nunziata, Ochel 2005), typically yielding inconclusive results. Pierre and Scarpetta (2004), while still relying on cross-sectional variation, use micro-data on firms. They show that firms in countries with more stringent employment regulations report being more hindered by these regulations, and that firms react to more stringent regulations by providing more training and resorting more to temporary employment. Although very valuable, such cross-sectional evidence may still be plagued by omitted variable biases, in as much as there are many unobservable country-specific factors that may be correlated with both firing regulations and firms' characteristics and behaviors.

It is thus important to examine the impact of variations in statutory firing costs within a single country. In recent years, several studies have used micro data to assess the consequences of changes in the regulation for one given country (e.g., Hunt 2000, Blanchard and Landier 2001, Kugler 2004, Kugler and Pica 2005). Most studies, whether cross-country or within countries, focus on the costs firms have to bear with certainty when firing under the regulations in place, setting aside the possibility of further intervention by labor courts. An exception is the study by Autor, Donohue and Schwab (2004) on the United States: using regional and temporal variation, they find a negative impact of one wrongful discharge doctrine, the implied-contract exception, on states' employment-to-population ratios. The implied-contract exception arises when, through words or actions, an employer implicitly promises not to terminate a worker without a good cause. Thus, the implied contract exception, a privately-granted right not to be unfairly dismissed, slightly reduces employment.

The timing of separations and the resulting duration of jobs have been subjected to both theoretical and empirical studies. A classic model by Jovanovic (1979) predicts a rise followed by a fall in the probability of separation with tenure. Productivity is job-specific and time-invariant; it is not known *ex ante* but becomes progressively evident as workers and firms observe output in succeeding periods. The probability of separation increases initially with the elapsed time because, as knowledge becomes more precise, the value of separating increases relative to the value of waiting to learn more about the real productivity of a job match whose current productivity is low. After some time, observed separation decreases because only the more productive matches remain. Farber (1994) empirically verifies Jovanovic's prediction about the relationship between tenure and separations. Using the National Longitudinal Survey of Youth, he shows that the monthly hazard of job separation initially increases with time spent on the job, peaks at 3 months, and decreases thereafter.

Here, I introduce three new elements of analysis. First, like Autor et al. (2004), I focus on labor courts' induced firing costs, and more specifically on the right not to be unfairly dismissed. But, instead of only examining the indirect effects of firing costs on employment, I directly analyze the effects of these costs on the probability of workers getting fired at different tenures. Second, I analyze the impact of firing costs on the timing, and not only the level, of firing. Third, I give this analysis a formal theoretical basis.

To test for the economic impact of a probationary period[3], I use the change in UK law mentioned above. Thus, the number of months necessary to qualify, or qualifying period, was lowered from 24 to 12 months for any termination (dismissal or redundancy) occurring after the 1st of June 1999. Employees with 12 to 23 months of tenure were not protected before the reform whereas they had the right to claim unfair dismissal if fired after the reform, implying that their probability of being fired should diminish after the reform. Employees with more than 24 months of tenure should be, in principle, relatively unaffected by the reform, and could be used as a control group. Employees with less than 12 months tenure may be affected by the reform if, for example, employers screen better after the reform to avoid a potential trial in the event of termination after the shorter qualifying period.

[3] While I am examining a tenure-dependant firing cost, another strand of literature examines the effect of a tenure-dependant quitting cost. Thus, in Canada, the tenure on the job necessary to qualify for unemployment benefits has varied, and a series of papers studies the effect of those changes on job duration (Baker and Rea, 1998, Christofides and McKenna, 1996).

The formal model I develop in Marinescu (2005) and summarize in section 3 gives further insights about the possible consequences of the reform on the firing hazard. The model's setup is very similar to Jovanovic's 1979 model, but some simplifying assumptions make it tractable, and suitable for calculating the impact of firing costs and other parameters on the hazard of firing. The model allows predicting how the hazard of firing should change after the reform if firms keep their personnel management policies fixed and only react to the shorter probationary period. The model also predicts how the firing hazard changes if firms react to the reform by increasing their recruitment or monitoring efforts, and it shows that these two strategies have significantly different effects. Thus, a higher recruitment effort implies a lower firing hazard for workers with 0 to a few months tenure, while a higher monitoring effort implies a higher firing hazard for these same workers.

The empirical analysis of the firing hazard uses duration models on the 2-quarters Labour Force Survey longitudinal datasets. A simple Kaplan-Meier estimate reveals that the firing hazard is indeed lower after the reform for employees with 12 to 24 months of tenure. The hazard is also found to be lower for employees with 0 to 12 months of tenure, which is consistent with firms having increased their recruitment efforts after the reform. Calibrating the model to fit these Kaplan-Meier estimates, I show that recruitment efforts must have indeed increased substantially after the reform, while monitoring on the job must also have increased slightly. Using all employees with more than 24 months of tenure as a control group in a Cox proportional hazard model, I find that the reform has a significant and large negative impact on the hazard of termination for those employees with 12 to 23 months tenure, and also for those with 0 to 11 months tenure. This result also holds if the control group is limited to employees with 26 to 48 months tenure. The estimated reduction in the firing hazard for workers with less than 2 years tenure relative to those with 2 to 4 years tenure is around 30%, with some small variation depending on the specification and the tenure sub-group considered. Lastly, I show that while most demographic and educational groups are similarly affected by the reform, the latter has a distinctive effect on university educated workers. After the reform, firms do not seem to increase recruitment efforts targeted at this latter group; instead, there is evidence consistent with a moderate increase in monitoring efforts.

I next look at the effects of the reform on wages, training, and the duration of unemployment. While this analysis is useful to better gauge the total impact of the reform on the economy, one should note two related caveats. First, the analysis lacks a firm theoretical basis as the theory developed in section 3 of this paper does not make direct predictions about these outcomes, and second, it is empirically weaker in as much as it is relatively hard to find reasonable control groups to identify the effects of interest. With these caveats in mind, results are as follows. First, no significant effect on wages can be established. Second, workers with 0 to 11 months tenure are significantly more likely to get training. The increase in training is consistent with an increase in match quality stemming from better recruitment and monitoring. Lastly, the reform was not associated with an increase in the duration of unemployment, but coincided instead with a decrease in unemployment duration for affected workers.

The rest of this paper is organized as follows. In Section 2, the tenure restriction to the right not to be unfairly dismissed is put into historical perspective. Section 3 presents the theoretical hypotheses to be tested, drawing on a model of learning about match quality. Section 4 describes the data, presents the main empirical results about the firing hazard, and analyzes the impact of the reform on the firing hazard of various sub-groups of the labor force. Section 5 analyzes the impact of the reform on wages, training and the duration of unemployment. Section 6 concludes.

## 2 The unfair dismissal qualifying period: historical background

The right not to be unfairly dismissed, introduced in most western European countries in the early 1970's, is usually restricted in several ways. One of the main restrictions is that employees must have a minimal period of continuous employment to fully qualify for this right[4]. In the UK, after Labour came to power in 1997, this qualifying period was lowered from 24 to 12 months by the 1999 Unfair Dismissal and Statement of Reasons for Dismissal (Variation of Qualifying Period) Order. This measure was part of a package destined to promote new labor practices. In the May 1998 *Fairness at Work* white paper (www.dti.gov.uk/er/fairness/), the New Labour government gave the following justification for the reduction in qualifying period:

> "As the economy becomes more dynamic, leading to more frequent job changes, the Government is concerned that this period is too long and a better balance between competitiveness and fairness would be achieved if it were reduced: employees would be less inhibited about changing jobs and thereby losing their protection, which should help to promote a more flexible labour market; more employers would see the case for introducing good employment practices, which should encourage a more committed and productive workforce. Some employers claim that a long qualification period is needed to allow mistakes made in recruitment to be rectified without

---

[4] For example, in France, while employees on unlimited term contracts (CDI) can always sue for unfair dismissal, they are only legally entitled to a minimum compensation for unfair dismissal if they have 2 or more years of tenure. This condition was set in 1973 when unfair dismissal legislation was first introduced, and has never been changed since. The introduction of the CNE contract in August 2005 could however be seen as an attempt to change this state of affairs since under that contract employees cannot sue their employer at all during the first 2 years of tenure, but the contract is identical to a CDI after two years of tenure. In the United Kingdom, the qualifying period is strict: employees cannot sue their employer for unfair dismissal if they have less than the minimum required tenure. Unlike France, the UK experimented a lot with the length of the qualifying period. Thus, while the initial 1971 (Industrial Relations Act) qualifying period had also been set to 24 months, it subsequently changed 7 times (Davies and Freedland 1993). Initially, all parties agreed to lower progressively the qualifying period so that all employees could be covered, and so by March 1975, the qualifying period had been reduced to 6 months. The main reason why the diminution in the qualifying period was to be progressive was that the newly created Industrial Tribunals could not immediately cope with a huge caseload.
However, by the end of the 1970's, and in particular after Mrs. Thatcher became prime minister in 1979, the terms of the debate changed. The right of employees to claim unfair dismissal was seen as a burden to businesses, in particular to small ones. By the time Mrs. Thatcher came to power, the qualifying period was down to 6 months. She immediately increased it to 12 months with the 1979 unfair dismissal (variation of qualifying period) order. Then the 1980 Employment Act increased this qualifying period again to 24 months for firms with less than 20 employees. Lastly, the 1985 "Unfair dismissal (variation of qualifying period)" order increased the qualifying period to 24 months for firms with more than 20 employees as well, which meant that by 1985 the qualifying period was 24 months for all employees.

> heavy costs. The Government accepts such mistakes happen but believes that the present period is longer than is needed to allow them to come to light and be dealt with. For all these reasons, and to increase protection against arbitrary dismissal, the Government therefore proposes to reduce the qualifying period to one year."

Thus, the reduction in the qualifying period is mainly seen as compensation offered to workers in exchange for their consent to a more flexible organization of the labor market.

Finally, one should note that the Labour government introduced a series of other labor market reforms that may potentially affect estimates of the impact of the change in the qualifying period for the right to claim unfair dismissal[5]. First, a National Minimum Wage was implemented in April 1999, and I will be correcting for this when relevant. Important new regulation has also been passed concerning parental leave and dependent care leave (Employment Relations Act 1999, and Maternity and Parental Leave Regulations 1999) and sex discrimination (Sex Discrimination (Gender Reassignment) Regulations 1999). These regulations mainly affect women, so it will be crucial to check whether estimated effects are driven by the female labor force. Lastly, the Employment Relations Act 1999 increased the limits on the awards workers who win a trial for unfair dismissal can get at court. However, the previous limit was already not binding: 95% of the awards workers obtained in 2003 (computed from the Survey of Employment Tribunal Applications, 2003, available on www.data-archive.ac.uk) were lower than the limit prevailing before 1999. It is therefore unlikely that this change has affected firms' behavior. Thus, while the regulatory activity had

[5] The right not to be unfairly dismissed is but one aspect of employment law regulating the termination of contracts of employment. Other important components are the notice period and the severance (or redundancy) pay rules. These latter features also depend on the tenure of the employee on the job, or more precisely continuous employment. The notice period is at least 1 week for more than 1 month and up to 2 years tenure, and at least 2 weeks for more than 2 years tenure, plus one additional week's notice for each further complete year of continuous employment for a period of less than 12 years' continuous employment; and at least 12 weeks' notice if the employee has been employed by the employer continuously for 12 years or more. Redundancy pay is only granted after two years of continuous employment and if the employee was fired for economic reasons. These features of employment law did not change in 1999, so it is important to bear in mind that the two years tenure may still be a meaningful juncture affecting firms' firing policies.

been intense at the time of the reform concerning the qualifying period for unfair dismissal, it seems feasible to identify its independent effects.

# 3 Model of the impact of firing costs on the timing of firing decisions

The right to claim unfair dismissal introduces a discontinuity in the cost of firing as a function of tenure on the job: when tenure becomes larger than the qualifying period, firing costs are suddenly augmented by the expected costs to the firm of possible unfair dismissal claims. The model I use is based on firm's learning about match quality, a hypothesis whose implications were first formally derived by Jovanovic(1979) and that was recently shown by Nagypal(2004) to be a driving factor of the empirical job separation hazard.

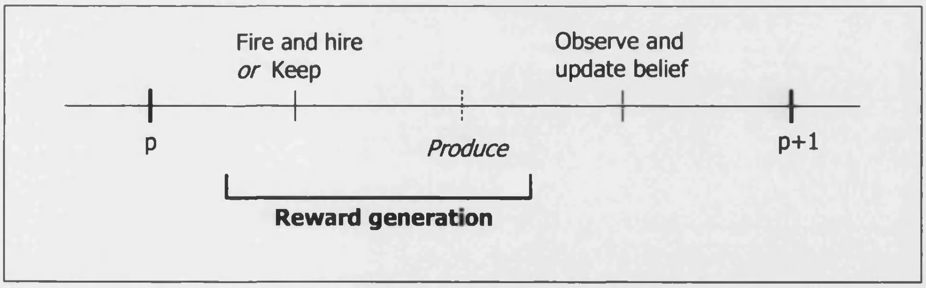
In what follows, I use a model based on dynamic programming developed in Marinescu (2005) to form testable hypotheses regarding the possible effects of a shortening of the qualifying period on the hazard of firing. The model's aim is to derive the firing hazard stemming from firms' optimal firing behavior in response to a set of parameters among which figures crucially the firing (and hiring) cost. The model necessarily involves many simplifications relative to actual firms' firing behavior. I defer a discussion of the model's limitations to Section 3.5.

## 3.1 Assumptions

When a firm and a worker begin their employment relationship they do not perfectly know their match quality[6] but learn about it over time. The worker is assumed to be passive in this model: the firm alone makes separation decisions.

The timing of events within each period p is formalized as follows:

[6] In what follows, I use the term "match quality", which given the literature usage suggests that match quality is idiosyncratic. However, as explained in section 3.5, I do not need for the purpose of this model to take a stance with respect to whether match quality is indeed idiosyncratic. Therefore, I could just as well use the term "worker quality" rather than "match quality".




The *set of possible actions* the firm can take is "fire the current worker and hire a new one", or "keep the current worker". Therefore, in this simple version of the model, unemployment or the overall level of labor demand are not modeled. Instead, the focus is on the efficiency and timing of the matching process.

The *state of the world* is defined by a vector of two variables: the tenure of the current worker, and the quality of the firm-worker match. The tenure variable is perfectly observed by the firm. Moreover, tenure cannot be higher than some tenure $t_{\max}$, which is to be conceived of as the retirement tenure. Match quality can be either good or bad[7]: a good match means that the worker is adequate for the job, whereas a bad match means that the worker is inadequate. I assume that a proportion $q$ of the matches is good whereas a proportion *1-q* is bad.

Match quality is not perfectly observed. Instead, at each period, the firm observes a normally distributed[8] signal about the quality of the match. The signal for a good match is normally distributed with mean 1 and variance $\sigma^2$, whereas for a bad match it is normally distributed with mean -1 and variance

[redacted]

normal distribution. Using Bayes' rule, one can then compute all possible beliefs $b(s,t)$ (see appendix 1 for the equation).

Using the Bellman equation, I can now specify the value as a function of the current belief. As in Jovanovic (1979), I assume that the firm only employs labor and has constant returns to scale. The actual per period return to a good match is 1 whereas the per period return to a bad match is 0. Moreover, the wage is fixed and set to 0[9]. Setting the wage to 0 rather than another constant does not entail any loss of generality given that labor demand is fixed in this economy and firms all pay the same wage. So if the firm keeps the worker, its expected return will be exactly $b(s,t)$. If the firm fires the worker, it gets the expected value of a new worker and incurs a separation (hiring and firing) cost $c(t)$ which is a function of the tenure $t$ of the current worker. I assume $c(t_{\max}) = c(1)$, i.e. when the worker retires, the firing cost is the same as the one incurred at tenure 1. This is because, at tenure 1 as at retirement, the separation cost consists mainly of the hiring cost of a new worker.

Let $V^*(b(s,t))$ be the value (i.e. the expected discounted future reward) of the match to the firm obtained when the firm follows the optimal policy.

The value of a worker to the firm if the firm keeps this worker (action $K$ ) is given by:

$$
\begin{aligned}
V(b(s,t),K) = \quad & b(s,t) + \\
& \delta.\{(1-b(s,t))\int_{-\infty}^{+\infty} f_b(s')*V^*(b(s',t+1))ds' + \\
& b(s,t)\int_{-\infty}^{+\infty} f_g(s')*V^*(b(s',t+1))ds'\}
\end{aligned}
\tag{1}
$$

The first line of equation 1 represents the immediate reward for keeping the worker, whereas the two following lines represent future rewards if keeping the worker at the current period, and are thus preceded by the discount factor $\delta$. The second line represents the future rewards if the match is bad

---

[9] One can also readily specify the wage to be a fixed share of the expected per period return, as would be the case with Nash bargaining. Qualitative results do not change when making this assumption.

weighted by the corresponding belief $1-b(s,t)$, whereas the third line represents the future rewards if the match is good weighted by the corresponding belief $b(s,t)$. For each of the two possible match qualities, the belief at the next period depends on the sum of signals $s'$ that the firm will have observed by tenure $t+1$, or equivalently on the signal at period $t+1$. Given my assumptions, if real match quality is bad and the sum of observations is $s$ (line 2 of equation 1), the probability of reaching a given $s'$ is given by a normal distribution $f_b$ with mean $s-1*(1-b(s,t))$ and variance $\sigma^2$ (remember that the mean of the per period signal for the low quality match is -1). A symmetric reasoning applies if the match is good and gives rise to line 3 of equation 1.

Alternatively, if the firm fires the worker (action $F$ ), the value is:

$$V(b(s,t),F) = V_{new} - c(t) \tag{2}$$

i.e. it is the value of a new worker minus the firing costs. Note that the value if fire only depends on the tenure due to the existence of tenure-dependent firing costs.

Given the values for keep and fire, the optimal value is given by the Bellman equation:

$$V^*(b(s,t)) = \max(V(b(s,t),K), V(b(s,t),F)) \tag{3}$$

Using dynamic programming and the appropriate Matlab code, the optimal policy of the firm is computed (see Marinescu(2006b) for the technical details). The policy can be expressed as a belief threshold $\tau(t)$ for each tenure $t$ such that if the firm's belief is equal to or above $\tau(t)$, then the firm keeps the worker, and otherwise it fires the worker.

The model so far has described the behavior of a representative firm. The behavior of infinitely many single-job firms can be represented by integrating the behavioral response of the firm over all the possible combinations of tenure $t$ and sum of signals $s$, given the assumed distributions. Thus, under the assumptions I use, it is possible to compute the firing hazard using the appropriate Matlab code. At tenure 1, the distribution of possible beliefs is computed given the assumed distributions. Then the hazard of firing at tenure 1 is the integral of the belief distribution from 0 to the firing threshold $\tau(1)$. At tenure 2, a set of possible signals is observed, which leads to a new distribution of possible beliefs, and the firing hazard is again the integral of the belief distribution from 0 to $\tau(2)$. And so on for each subsequent tenure (see Appendix 1 for the equation). Note that the computation does not rely on simulation, i.e. the technique used does not involve drawing a large number of matches in conformity with the distribution and then averaging over the results. Instead, equations directly use the definitions of probability distributions, and computations rely on an approximation of the normal distribution of the signal by a finite number of points.

## 3.2 Parameters

I now proceed to examine the effects on the hazard rate of termination of a discontinuity in firing costs (with higher firing costs after a given tenure) and how the hazard rate changes when the length of the probationary period changes. I thus model the potential effects of the 1999 reform within the framework of this model.

I choose a benchmark case for clarity of exposition. The parameters were chosen so that the shape of the hazard curve is similar to the hazard of firing observed in the United Kingdom in 1996-1999 (shown in Figure 8). Moreover, in this benchmark case, I pay attention to choosing parameters so that the variations in these parameters show sufficiently large effects to be clearly visible on graphs. When analyzing actual data, I will directly fit the theoretical hazard curve to the empirical one and derive the underlying parameters.

The parameters of the benchmark case are as displayed in Table 1. A firing (and hiring) cost of 7 corresponds to 7 months of output. Note that an increase in the maximal tenure does not change the hazard of firing for tenures 1 to 50, tenures on which I will be focusing.

I introduce a tenure-dependant firing cost in the following form. The firing cost is 7 before the end of the probationary period, and 9 thereafter. I start with analyzing the effects of different lengths of the probationary period.

## 3.3 Variation in length of the probationary period

The hazard of firing is determined by two factors: the firing threshold $\tau(t)$ expressed in terms of belief, and the distribution of the firm's belief. The latter distribution is itself determined by two factors: the distribution of match quality embodied in the $q$ parameter giving the proportion of good matches, and the distribution of signals engendered by the variance $\sigma^2$. Let us first consider the case where the firing cost does not vary with tenure but is instead fixed at 7. The firing hazard is plotted in Figure 1. It is first increasing and then decreasing in tenure, as in Jovanovic(1979). In Figure 2, I plot the distribution of firms' beliefs at different tenures, after they have observed the signal at that tenure and before they fire. First, note that at tenure 1, right before firms have their first opportunity to fire, the distribution of beliefs about match quality is roughly normal with a mean of 0.5, corresponding to the $q$ I specified. Now, up to tenure 48[10], the firing threshold is constant at .22, i.e. if the probability that the match is good is 22% or more, the firm keeps the worker, and otherwise it fires. The firing range of belief is shaded in Figure 2. As already mentioned, the firing hazard is the integral of the belief distribution below the threshold[11], i.e. in the shaded area. This explains why, as tenure increases, the distribution moves away from normality: indeed, as firms fire the worst-performing workers, they truncate the lower tail of the distribution. Thus, as tenure increases, the distribution of belief in the neighborhood of .5 flattens and its mean moves towards 1: this is because in the long run, firms only keep workers whose match quality is almost certainly high. The shape of the belief distribution by tenure also helps to understand why the firing hazard

[10] The firing threshold changes slightly thereafter because the firm anticipates that the worker is going to retire after 200 months.

[11] The reader may have noticed that these integrals at different tenures do not perfectly square with the firing hazard plotted in Figure 1. This is because the computation of the hazard is based on the "sum of observations" statistic. While this latter statistic translates unequivocally into a given belief (see the formula for the belief in appendix 1), converting a distribution in terms of sum of observations to a distribution in terms of belief entails a certain degree of approximation because of the discretization used for the "sum of observations" statistic.

first increases and then decreases with tenure. Indeed, if firms never fired anyone, the belief distribution would be more and more concentrated at 0 and 1 with increasing tenure. This is because as firms observe more and more signals, they improve their inference about whether a match is good or bad, and thus after an infinite number of observations, the belief distribution would be two-peaked with a density of .5 (because of the parameter q=.5) at 0, .5 at 1, and 0 everywhere else. Thus, in the absence of firing there would be more and more workers below the firing threshold with increasing tenure, so that the potential firing hazard would monotonically increase with tenure. This explains why the firing hazard first increases with tenure: more and more matches are discovered to be of bad quality as tenure increases. But as firms always dissolve the worst quality matches, eventually a large proportion of matches will actually be good and so there will be very few workers for whom the belief can fall below the firing threshold. This is why the firing hazard eventually decreases.

What is the effect of the introduction of a probationary period? To illustrate this effect, I assume that at tenure 24, the firing cost goes from 7 to 9. This only affects the firing hazard through the threshold, and not through the parameters determining the belief distributions since $\sigma^2$ and $q$ remain unchanged by assumption. With a higher firing cost after 24 months, the threshold will obviously decrease for tenures greater than 24 months, i.e. as firing is more expensive, firms keep workers with lower believed match quality. So the threshold after the end of the probationary period will be lower with a probationary period than without. What happens to the threshold before the end of the probationary period? First, at low tenure, the threshold for firing is the same as in the absence of a probationary period. This means that the hazard will also be exactly the same at low tenure, as seen in Figure 1. Then, as tenure increases, firms anticipate that there will be a higher firing cost in the near future, so they *increase* their threshold before the end of the probationary period, thus firing preventively a group of workers whose match quality is fairly low and who would otherwise be likely to get fired at higher cost after the end of the probationary period. This is what creates the spike and the trough in the firing hazard with 24-months probationary period seen in Figure

1: indeed, right before the end of the probationary period, more workers get fired because of the higher firing threshold, whereas right after the end of the probationary period, less workers get fired because the threshold is lower *and* those who were most likely to fall below it have been fired preventively. While it is the case that the firing threshold is lower in the post-probationary period, the firing hazard at tenures higher than 35 is almost the same as in the absence of a probationary period: this is because at that point very few workers get fired, for example when one looks at the distribution of beliefs at tenure 35 in Figure 2, one can see that the area between .22 and .2 (that is, the threshold at tenure 35 with a 24 months probationary period) is fairly small, and so moving the threshold down to .2 has a relatively small effect compared to, say, the same downward move of the threshold occurring at tenure 5.

What is the effect of a shortening of the probationary period? The firing cost is assumed to increase from 7 to 9 at tenure 12. This implies that the firing threshold will decrease earlier due to higher firing costs setting in earlier, and so the increase in the firing threshold before the probationary period will also occur earlier. For the shape of the firing hazard, this implies that while the firing hazard will remain exactly the same at very low tenure, the spike and trough will occur earlier, while there will be little effect on the firing hazard at high tenures, which is what can be seen in Figure 1.

This analysis however does not take into account the fact that firms could be endogenously reacting to the shortening of the qualifying period by increasing the quality $q$ of matches when hiring, or by increasing the intensity of monitoring on the job and thus decreasing $\sigma^2$. Intuitively, both strategies would reduce the probability that firms should have to fire after the end of the probationary period. That this can be an optimal reaction on the part of firms is further confirmed by the following consideration. In computations not reproduced here, I found that, starting from the reference case, the marginal gain (as measured by the change in the value of a new worker) of increasing either recruitment or monitoring intensity is greater with a shorter probationary period. This implies that, for a given marginal cost of these technologies, firms should be more willing to invest in them after the reform. Moreover, increasing the recruitment intensity yields a higher marginal gain than increasing the monitoring intensity. Thus, if the marginal cost of recruitment effort is not much greater than the marginal cost of monitoring effort, we expect firms to increase recruitment intensity more than monitoring intensity after the reform.

## 3.4 Endogenous response: modification of the quality of recruitment or monitoring

I study here the effects on the firing hazard of increasing the recruitment quality $q$ from .5 to .7 or increasing the monitoring intensity, i.e. decreasing $\sigma^2$ from 16 to 4. The corresponding curves are plotted in Figure 3.

An increase in recruitment quality results in a decrease in firing at all tenures. This effect can be decomposed in two elements (which are in fact jointly determined, and only separated for the purpose of exposition). First, the increase in $q$ increases the firing threshold from .22 to .34 in the 8 first months of tenure, which, for the belief distributions with $q$=.5, would imply more firing. But second, the increase in $q$ changes the shape of the belief distribution by tenure as shown in Figure 4, i.e. it changes how *likely* it is that the firm holds a belief below the threshold. Figure 4 shows that the means of the distributions are shifted rightwards, so that the lower tails of the belief distributions are thinner, which implies less firing. To understand why, as can be seen from the hazard curves in Figure 3, the effect on the belief distribution dominates the effect on the threshold, we must consider the following. First, note that the firm's belief exclusively depends on the sum of observations for a worker. Therefore, the threshold may also be equivalently expressed in terms of sum of observations. The threshold expressed in terms of sum of observations goes from -10 to -12 when recruitment effort increases (remember that a good match generates on average an observation of 1 per period whereas a bad match generates an observation of -1). In other terms, firms wait for more negative observations before they fire someone. This is intuitive because now they have a higher prior: they know that 70% instead of 50% of matches are good. Therefore any bad observation is more likely to be just noise. The fact that the threshold goes down in terms of sum of observations also implies that it is less likely that someone gets fired in general, because it is less likely that the sum of observations be below -12 rather than below -10: indeed with half good matches and half bad matches as in the reference case, the average observation will be 0; moreover, with an increase in recruitment efforts there are not 50% but 70% of good matches so it is even less likely that the sum of observations for a worker shall fall below -12. Therefore, the hazard of firing should fall. So why is the threshold expressed in terms of belief higher? This is because while firms wait for more negative observations before they fire someone, at the same time they know that there are more good matches in the population of potential employees: this therefore makes them slightly more demanding on the current employees.

By contrast, an increase in monitoring results in an increase in firing at low tenures and a decrease in firing at high tenures (Figure 3). This results again from two effects. First, the firing threshold decreases from .22 to .12 in the first 8 months, which for the belief distributions with $\sigma^2 = 16$, would imply less firing. Second, the shape of the belief distributions changes, as shown in Figure 5: the distributions are flatter than before in the neighborhood of .5. To understand why this is the case, let's take the distribution at tenure 1, before the firm has had any chance to fire. This distribution is flatter because the signals are more informative than before: so, instead of having the belief distribution highly concentrated around .5, which is the prior over the population of hired workers, the belief distribution has more weight on its tails, because even after one signal firms are already quite certain that some matches are bad while others are good. This change in the shape of the distribution entails more firing at low tenures, because now for any threshold below .5, there are more workers below this threshold at low tenures. But eventually, because firms can quickly get rid of bad matches, the hazard of firing gets lower. In this case, when expressing the threshold in terms of sum of observations, this threshold goes up to from -10 to -4. With an unchanged proportion of good matches (50%), it is more likely that a random worker has a sum of observations below -4 instead of below -10. Therefore, at low tenure, when firms did not yet get to fire many people so that the population of employed workers is still similar to the population of employable workers, it is more likely that someone gets fired. The firing hazard is thus higher at low tenure. So then why does the belief threshold go down? This is because as firms get more precise information every period, they can afford to wait a little bit longer to be really sure that a match is indeed likely to be bad and therefore worth terminating.

It is also possible to compute the value of a new match under higher recruitment or monitoring efforts, and compare it to the value in the reference case. Thus, if the firm increases recruitment or monitoring effort at no cost, then all other things equal the value of a new match increases. With higher recruitment efforts, by definition the firm gets better matches on average from day one of the employment relationship, whereas with higher monitoring efforts the firm eventually gets better matches as it is better able to detect and dissolve the bad matches once the employment relationship has begun. Moreover, both increasing the recruitment effort and the monitoring intensity indeed decrease the hazard of firing after the probationary period but they have opposite effects on firing at low tenure (i.e. for tenures between 0 and a few months) : while an increase in recruitment effort decreases firing at low tenure, an increase in monitoring increases it.

## 3.5 Limits to the model

The first limit to the model developed above is that match quality can only take two values, good or bad. However, in Marinescu (2005), I show that the qualitative implications of the model are preserved if one uses a more continuous distribution, such as for example a normal distribution. Second, there is no explicit cost for the firm of increasing recruitment efforts or monitoring. In the absence of recruitment and monitoring costs, firms could completely change the parameters so that uncertainty would no longer be a problem, and they would only get good matches. In reality, these efforts are of course costly and the reduction in uncertainty and increase in match quality will only be obtained if cost-effective. Note however that the costs of these efforts can be viewed as part of the separation cost if assumed to be a fixed cost per match.

A more important limitation of the model is that it relies on partial equilibrium analysis. Thus, I am not modeling the influence of the behavior of one firm on other firms' behavior, nor the aggregate demand for labor. Therefore, I do not need to take a stance with respect to whether match quality is in fact idiosyncratic (Jovanovic 1979) or whether there is some symmetric (Gibbons, Katz, Lemieux, and Parent 2005, Moffitt and Jovanovic 1990) or asymmetric learning about general ability (Gibbons and Katz 1991, Schoneberg 2004). Nevertheless, the nature of the information imperfection about match quality may have important effects when evaluating the overall efficiency and welfare effects of a change in firing costs. For example, if firing costs get higher and there is asymmetric learning about quality, then all else equal, the believed average quality of terminated matches diminishes, implying that terminated workers have lower reemployment probabilities. However, in this model I am focusing on what drives firms' firing behavior, and it is only when looking at other outcomes such as unemployment duration in the empirical analysis that I will briefly consider the implications of different possible hypotheses about match quality.

## 3.6 Main conclusions drawn from the model

The main conclusions drawn from the model are summarized in Table 2. Note that it is not possible to determine in the general case what happens for workers who have tenures just below 12 months: indeed, the shortening in the probationary period implies that there should be a spike before 12 months, but if other parameters such as $q$ or $\sigma^2$ change then this spike may lie below the curve corresponding to a 24 month probationary period. For the purpose of empirical analysis, the most important lesson from the theory is that it is by looking at workers with low tenure that one can hope to distinguish among the different scenarios summarized in Table 2. It is moreover important to note that while the absolute size of the effects of large changes in recruitment and monitoring efforts on the hazard of firing for workers with 0 to 24 months tenure is large, effects are very limited for workers with more than 24 months tenure (see Figure 3). This implies that, form a theoretical perspective, workers with more than 24 months tenure should form a reasonable if imperfect control group.

# 4 The impact of the reform on the firing hazard

Before moving on to the description of the micro dataset used in this paper, it is useful to first have an idea of the macroeconomic context in which the reform takes place. I thus plot in Figure 6 the evolution of the employment-to-population ratio in the United Kingdom in the long run. The focus of this paper, the 1999 reform, occurs during a phase of steadily growing employment in the UK, and the reform does not have any immediate impact on the growing employment trend. While employment growth does slow down from August 2000 onwards, it is difficult to attribute this to the reform. By the beginning of 2005, the employment to population ratio reaches an almost all time high; it is only surpassed by the values observed before 1976. Thus, it is unlikely that the 1999 reform has had any major impact on average labor demand in the British economy.

## 4.1 Data

The British Labour Force Survey (LFS) is administrated each quarter and contains questions similar to the Current Population Survey in the US. It covers women from 15 to 59 years old, and men from 15 to 64 years old at the date of the first interview. It is a rotating panel, and each household[12] remains in the sample for 5 months. This paper uses the 2-quarters Labour Force survey longitudinal datasets[13] from March 1996 to September 2004. These datasets are put together by the UK Office of National Statistics and they contain all occurrences of individuals in the LFS being observed in two consecutive quarters.

The right to claim unfair dismissal only applies to employees (i.e. not self-employed, of course!) in permanent jobs working usually more than 16 hours a week. I therefore restrict my main sample[14] to those employees. In principle, workers on fixed-term contracts also have the right to claim unfair dismissal, but before 1999 (Employment Relations Act), they could contractually waive this right. Moreover, the majority of employees on fixed

---

[12] Households in the sample are identified by their addresses so people who move during the survey drop out of the sample.
[13] Full documentation about the datasets can be found on www.data-archive.ac.uk.
[14] A different sample will be used to study the duration of unemployment.

term contracts have a tenure inferior to 2 years, which makes identifying the probability of being fired after 2 years difficult. Altogether, this means that analyzing the effects of the reform for this group would not be as instructive as for permanent workers. I therefore perform the analysis on the latter group only[15].

Because the dataset is a panel, a job can be observed for two or more consecutive periods. I only keep the first observation for each job. Thus several jobs held by the same person can be present in the sample, but not the same job observed at two or more different points in time. When it is possible, I will therefore cluster by person, and when not I will only keep the first job observed for each person.

Having defined the relevant group of workers, I also have to compute their tenure. The date of hiring is present for more than 99% of currently employed workers along with the date of the interview. In most cases, both the year and month of hiring are known, which allows for the computation of the tenure in months. When only the year of hiring is known, and the worker has less than 4 years tenure, I drop the observation because monthly precision is important in that range; otherwise I keep it and assume the month of hiring was January (this is random with respect to each job). For workers who separate from their jobs, the tenure at separation can also be calculated. For those who are still unemployed by the second quarter, the date when their last job ended is present. If however workers have found a new job, the date when they left their last job is not present, so it has to be imputed. The distribution of completed unemployment spells lasting 3 months or less and beginning and ending with employment has 3 months as a mode. Therefore, I assume that if a worker separated from the job he was holding in the first quarter and found a new job by the second quarter, then he separated from the first job during the month of the first interview, i.e. I make the unemployment spell as long as possible in order to conform with the distribution of completed unemployment spells. Using the hiring date workers provided in the first quarter of

[15] I performed the analysis of the impact of the reform on employees on temporary jobs, i.e. fixed term contracts, seasonal work and agencies, and found that there is no impact of the reform (results not reproduced here).

observation and the date when they left their job or the imputation thereof, I can thus compute their tenure in months at the moment of termination.

What are the potential tenure sampling problems? The sample of jobs is what is traditionally called in the duration literature a stock sample with follow-up: one observes the tenure of workers in employment at the date of the first interview (stock sample), and then whether they separate by the second interview (follow-up). This causes two problems. First, long tenures are overrepresented in the sense that one observes a higher proportion of high tenure workers in the sample than would be observed in a flow sample, i.e. in a sample where one can follow workers from day one of their job. Indeed, all the jobs that started x years before the first period of observation and ended in the meantime are not observed. However, it is possible to correct for this bias in survival analysis by specifying the date of entry in the study, which in this case will be the date of the first interview[16]. Second, the follow-up also causes a small problem if a job begun *and* ended during the 3-months period between two interviews. In that case, I make a wrong inference about which job was left and when: indeed, I will be assuming that the job left by the second quarter was the job observed at the first quarter, whereas in fact it was another short job that followed in the meantime. To document the prevalence of such a problem, I compare the characteristics in terms of occupation and industry of the *last* job held as described in the second quarter interview with those of the job that was held in the first quarter. As it happens, when the information on both jobs is available, there is a discrepancy in only 4% of the cases, and I decide to drop these latter cases.

If a worker left his job in the previous quarter, he is prompted to indicate the reason why the job ended among a list of the following possibilities: dismissed, made redundant, temporary job finished, resigned, gave up for health reasons, took early retirement, retired, gave up for family or personal reasons, other reason. When using duration models to explain a given type of

[16] To be precise, I use as date of entry in the study the date of interview minus one month. This is because Stata drops all observations for which a failure is observed at the date of entry, and I just mentioned how some workers either lost their job during the month of their first interview or otherwise were assumed to have done so. The whole small subtlety occurs because we have discrete time steps that are long enough (one month) to contain both the interview and the job loss.

separation, I treat other types of separations as censoring. In this section, I mainly focus on workers who were fired, i.e. dismissed or made redundant, since they are the ones directly affected by the law.

To summarize, the main sample consists of employees in permanent jobs usually working more than 16 hours per week and having a known tenure. Table 3 gives summary statistics for the sample used. Note that among the reasons given by workers for leaving their last job, dismissals and redundancies represent a sizeable 21.7%, a proportion comparable to the "other" category (22.4%) but lower that quits (35.6%). Since the question involves self-reporting, the distinction between dismissals and redundancies has to be taken with skepticism: indeed, workers may prefer to report that they were laid off rather than discharged. It is somewhat puzzling that the end of a temporary job is a reason quoted by 3.4% of workers although the sample includes permanent jobs only; however, while the question asking about permanent jobs prompts the worker to clearly indicate if the job is *objectively* temporary rather than temporary because he intends it to be temporary, this distinction is not insisted upon in the question about the reason for leaving the last job. Therefore, it could be that these workers meant that that job was temporary for them.

I now focus on workers who were fired.

## 4.2 A first look at firing rates by tenure

Assuming, consistent with the model, that workers with more than 24 months tenure are a reasonable control group, I plot the raw monthly job loss rate by tenure range in Figure 7. The raw job loss rate is defined as the number of employees who lost their job through dismissal or redundancy over the total number of employees in the sample. Although there is a lot of month-to-month variation, one observes that globally the job loss rate of the control group (the more than 24 months tenure) is stable during the period observed, with some minor decrease in mid-2001, and some slightly higher values after the world economic downturn following September 2001. On the other hand, the treated group, i.e. the employees with less than 24 months tenure, has a decreasing trend in its firing rate starting after the June 1999 reform, so that at the end of the observation period the job loss rates for the treated group are smaller on average than at the beginning of the observation period, and they are also almost undistinguishable from the job loss rates of the control group.

This preliminary graphical analysis thus seems to indicate that the job loss probability of the treated group is negatively affected by the reform. In other terms, the reform seems to have decreased the separation probability for employees with less than 2 years tenure. I now investigate how the reform affected the hazard of firing for all tenures.

## 4.3 A Kaplan-Meier estimate of the hazard of firing

I plot the non-parametric Kaplan-Meier estimate of the hazard of firing before and after June 1999 (Figure 8). Like Farber(1994), I find a pattern consistent with Jovanovic's 1979 model, and the model developed in section 3. While the peak in terminations occurs at about 3 months as in Farber's work, it is not as sharp. This difference is not due to my looking only at terminations and not at quits, as performing the same analysis on quits yields a similar pattern (see Appendix 2 Figure 12). It is instead likely to be due to the fact that the NSLY is a sample of young people. Indeed, I find that for people aged less than 40, there is a sharper peak at 3-4 months. The model developed in section 3 and in Marinescu (2005) suggests that the observed difference between younger and older workers' firing hazard can be explained by higher firing and hiring costs for older workers, or by a greater per period uncertainty about older workers' performance.

Figure 8 shows that the shape of the hazard function in the before period is very similar to the theoretical hazard curve corresponding to a 24 months probationary period in Figure 3: in particular, one very clearly observes a trough in the firing hazard around 24 months. With respect to the change introduced by the reform, one observes that from 24 months on, the hazard function is essentially identical before and after the reform. This confirms that employees with more than 24 months of tenure form a good control group. The hazard of termination after the reform is significantly lower on the interval [0,24]. It is thus lower not only on the interval [12,23], but also on the interval [0,12], which indicates that it is likely that the quality of recruitment has increased (see model's predictions in Table 2). Note that while there is no observable change in the firing hazard for the 24 to 48 months tenure group, this does not contradict the model's predictions in the case of an increase in recruitment effort. Indeed, the decrease in the firing hazard for the 24 to 48 months tenure group engendered by an increase in recruitment quality is likely to be very small (see section 3 and Figure 3).

In order to estimate how big a role increases in recruitment and/or monitoring efforts play in explaining the change in the shape of the hazard function after the reform, it is informative to perform a model calibration exercise: what are the parameters of the model that best correspond to the Kaplan-Meier empirical hazard before and after the reform? While imperfect due to the limitations of the model and the calibration procedure, this exercise is useful to build quantitative intuition about the effects of the reform on the firing hazard. The calibration procedure looks for the parameters of the model that minimize the sum of the squared differences between the theoretical and the empirically estimated firing hazard curves[17]. The fixed parameters in the model are the same as in section 3. The results of the calibration exercise are shown in Table 4. I begin with fitting the hazard in the pre-reform period. I find that the best fit implies a total firing and hiring cost of 6.6 during the first 24 months, and 6.8 thereafter. To judge how big these costs are, the reader is reminded that a good match produces a value of 1 per month. Thus, firing and hiring costs are somewhat higher than 6 months of output. The proportion of good matches is 41%, and the standard error of the observation is 5.7. The calibration thus implies that the matching technology is not too efficient and that firing and hiring costs are high.

Now, what about the impact of the reform on those parameters? I first set the length of the probationary period to 12 months, and I look for the best $q$

[17] It uses the Matlab function fminsearch to do so. Note moreover that I decide to calibrate the model to best fit the 36 first months of the empirical hazard function in the case of a 24 months probationary period, and the 24 first months of the empirical hazard function in the case of a 12 months probationary period. The model is indeed inadequate at explaining firing hazards at high tenure for structural reasons, and so imposing that the model should fit the firing hazard at high tenure uselessly damages the quality of the fit at low tenure. Indeed, the theoretical firing hazard decreases very fast to 0 for high tenures, as almost all bad matches have been dissolved, whereas the empirical hazard remains roughly at the same level beyond 30 months of tenure. This is very likely due to the fact that match quality is not in reality constant over time, as assumed by the model, but good matches may turn bad (see Marinescu(2006b) for a model that includes this feature).

and $\sigma$ parameters to fit the post-reform empirical firing hazard, leaving all other parameters as before. I find that the quality of recruitment has increased a lot from 42% of good matches to 63% of good matches. Moreover, monitoring intensity must also have slightly increased as the standard error decreased. In the third column, I use an alternative calibration procedure where I also allow the firing and hiring cost during the probationary period to vary. The reason for doing so is that if the recruitment efforts have increased, then hiring costs must have increased as well. The calibration results shown in the third column imply a substantially higher firing and hiring cost during the probationary period: indeed, the latter is now almost as high as the cost incurred after the probationary period (6.782 versus 6.8).

The calibration thus confirms the increase in recruitment effort in the post-reform period – an inference which could already be made by observing the empirical hazard function and using the model's predictions – and quantifies that increase. The calibration also shows a small increase in monitoring intensity, which could not be inferred by looking at the shape of the empirical hazard function but is consistent with what could have been expected ex ante. Thus, the reform seems to have encouraged firms to increase the quality of their recruitment and the intensity of monitoring. One way to check for the plausibility of this prediction is to rely on the fact that increasing monitoring or recruitment effort is likely to take some time while reducing the firing hazard of workers with 12 to 23 months tenure can be done more quickly. In this case, over time, one should observe that the hazard of firing first diminishes for the 12 to 23 months tenure, and then for the 0 to 11. This is indeed what I find when I plot the hazards using one year of data at a time (results not reproduced here). The way the hazard of firing changes through time is thus consistent with firms first directly reacting to the reform by firing less workers with 12 to 23 month tenure, and then increasing recruitment and monitoring efforts.

Another way of checking for the plausibility of the model's predictions is to look for other evidence about firms' recruitment and monitoring practices. One such piece of evidence is the 2004 Workplace Employment Relations Survey (WERS 2004). Data from this survey has not been made available yet, but I can draw on a summary of results by Kersley et al. (2005).

Between 1998 and 2004, there has been no substantial change in the use of tests by employers when recruiting employees. Thus, if recruitment efforts are measured as the use of tests, there does not seem to be a substantial increase in recruitment efforts. However, this measure of recruitment efforts seems overly restrictive. Consistent with an increase in monitoring, performance appraisals are more widely used after the reform: while 73% of employers used them in 1998, 78% did so in 2004. Another source of evidence on employers' reaction to the qualifying period for unfair dismissal is the Blackburn and Hart (2002) report on small firms' (i.e. with more than one but less than 50 employees) awareness and knowledge of individual employment rights. Employers report that unfair dismissal is the most constraining regulation after the minimum wage and maternity rights. In July-August 2000, 65% of these small employers were aware that there exists a length of service necessary to qualify for unfair dismissal, but their estimates varied between 1 week and 3 years, with a mean at 15 months, which is somewhat higher than the qualifying period prevailing in 2000. Lastly, employers also reported that because of the risk of an unfair dismissal trial, they are taking more care about who they recruit, which is consistent with an increase in recruitment efforts.

Having thus examined the basic patterns of change in the firing hazard, I move on to a more systematic approach, controlling for other variables that may have affected the hazard of firing.

## 4.4 Controlling for covariates using a Cox proportional hazard model

To test the robustness of my findings, I estimate a Cox proportional hazard model with delayed entry[18], controlling for essential covariates. The advantage of such a model is that there is no need to specify the functional form of the baseline hazard (Lancaster 1990).

To test for the effect of the 1999 reform, I use two related procedures. First I plot the baseline hazard of firing before and after the reform in Figure 9. The method used here is to run a stratified Cox model and compute the

[18] As explained in section 4.1, jobs are at risk of being terminated from the date of hiring but they are only observed from the date of the first interview on, i.e. they enter the study with a delay.

baseline hazards for the strata "before" and for the strata "after". The stratified Cox model assumes that the coefficients on the control variables are the same before and after the reform. Figure 9 is almost identical to the Kaplan-Meier plot in Figure 8 implying that controls for covariates do not change the main conclusions.

I then proceed to run a Cox regression with the following specification for the hazard of termination:

$$\lambda(t,Z) = \lambda_0(t)\exp\{\beta' Z(t) + \gamma_0' Treat + \gamma_1' Treat * After\}$$

$Z$ is a set of controls, including a full set of year dummies. *Treat* is a set of dummies for different ranges of tenure within the treatment group, i.e. employees with less than 25 months of tenure. *After* is a dummy that takes the value one from June 1999 on (or that takes the value 1 from June 2000 on and is missing from June 1998 to May 2000, depending on specifications). *Treat*After* is the interaction between *Treat* and *After*. The *Treat* dummies measure how the hazard of termination for the treatment group systematically differs from the hazard of termination for the control group. A test of the negative effect of the reform on the hazard of termination is that the coefficients in the $\gamma_1'$ vector are negative and significant.

Panel A of Table 5 presents the results using basic tenure categories for the treated groups, that is 0 to 11 months and 12 to 23 months. Using After 1999 as the reform dummy, I find that the reform significantly reduced the firing hazard by 18% for workers with 0 to 11 months tenure and by 20% for workers with 12 to 23 months tenure relative to those workers having more than 24 months tenure. I can also use as a control the workers with 24 to 48 months tenure, as they are likely to be more similar to the 0 to 23 months tenure group than workers who have tenures above 48 months. Using this control group does not change the results: if anything, the effect of the reform is now stronger. A problem with using "after June 1999" as the post-reform period is that firms may have anticipated the reform and/or it may have taken some time for firms to adjust to the new regulation. Therefore, I use as an alternative measure the after period "after June 1999, but excluding observations from May 1998 to May 2000". The results are not, however, affected by this change in the definition of the reform period[19].

In panel B, I use detailed tenure categories to examine the effects on different tenure subgroups. Again, the choice of control group or post-reform period does not change the results. I therefore concentrate on the more demanding specification, i.e. taking the 24 to 48 months group as a control and using "after June 1999, excluding May 1998 to May 2000" as the post-reform period. This is also the specification I adopt in the rest of the paper, unless otherwise specified. Concerning the effect of the reform on different tenure categories, I find that the negative effect of the reform on the firing hazard is significant for all subgroups up to month 21, and fades away from month 22 to months 25. The effect is of similar magnitude as in panel A, implying a reduction in the firing hazard of about 30 to 40% for all subgroups from month 5 to month 21, with a somewhat smaller effect for the 0 to 4 months tenure group. The fact that the effect is smaller for that very low tenure group was to be expected from the observation of Figure 3 (compare the "24 months prob. period" curve with the "12 months prob. period, q=.7" curve) and Figure 8. The reduction in the firing hazard is largest for the 18 to 21 months tenure, likely due to the fact that before the reform there used to be a spike at about 21 months tenure (Figure 8).

In general, the reform is found to be effective in lowering the hazard of firing for the group newly protected by the right to claim unfair dismissal, i.e. the 12 to 23 months tenure group. Moreover, it also significantly lowers the hazard of termination for workers with 0 to 11 months tenure, which is consistent with the employers having increased their recruitment efforts in reaction to the reform.

## 4.5 Impact on different groups

In this section, I test whether the reform has heterogeneous effects on subgroups of workers. Indeed, numerous papers studying the impact of firing

---

[19] I also used two other definitions of the reform dummy. In one case, I only allowed for an anticipation effect, excluding the period May 1998 to June 1999, and in the other I only allowed for an adaptation effect by excluding June 1999 to May 2000. The results in presented in Table 5 are however unaffected by these alternative definitions.

costs found that higher firing costs tended to mostly protect prime-age males and more educated workers while negatively affecting youths, females and the less educated (see for example OECD(1999) or Blanchard and Landier(2002)). It is therefore interesting to ask if this tightening in workers' protection against unfair dismissal affected differentially these latter groups. Moreover, analyzing the effects by sub-groups allows for better estimation in as much as fewer constraints on the parameters have to be imposed. Indeed, different worker types have different underlying parameters affecting the shape of their firing hazard, while the Cox specification only allows for proportional shift with covariates. If a sub-group has an altogether different pattern of firing hazard by tenure, then the Cox specification does not properly take that into account. This means that it is useful for identification purposes to separate the sample in more homogenous sub-groups.

Table 6 examines the effects of the reform by gender and age, while Table 7 looks at education. Panel A of Table 6 shows the break-down by gender. While females see a somewhat higher decrease in their firing hazard than men, this difference is not significant. Thus, reforms in the areas of dependent care and sex discrimination, which intervened at the same time as the reform of interest, are not driving the results. Panel B shows the break-down by age. The effect on the 0 to 11 months tenure group is basically the same for old and young workers, whereas the effect for 12 to 23 months tenure group is more pronounced for younger workers.

Table 7 shows the impact of the reform on the firing hazard by level of education. The hazard of firing significantly decreases for workers with 0 to 23 months tenure who are less than college educated, but not for those who are college educated. For workers with 12 to 23 months tenure, the hazard of firing decreases for all levels of education, even though the point estimate of the decrease in the firing hazard for university educated workers with 12 to 23 months tenure is lower and insignificant. Why are university educated workers different? When looking at the Kaplan-Meier plot of their hazard of firing before and after the reform (figure not reproduced here), it appears that the positive insignificant effect of the reform (Table 7) on workers with 0 to 11 months tenure is due to the fact that after the reform the peak in the firing hazard occurs at 7 months, while it occurred at 12 months before the reform. Moreover, while the trough in the firing hazard at 24 months was much bigger for university educated workers than for the whole population before the reform, it completely disappears after the reform. The model in section 3 explains these results. First, the peak in the firing hazard occurs later for higher educated workers than for others because these workers are more costly to fire and hire and/or harder to monitor. Both assumptions seem realistic in the case of university educated workers. However, after the reform, firms can no longer wait so long before they fire because with the new 12-months probationary period they would incur too high a firing cost; thus the peak in the firing hazard occurs before 12 months after the reform, consistent with an increase in monitoring effort. Moreover, the model tells us that the hazard of firing will only decrease at low tenures, i.e. here for workers with 0 to a few months tenure, if the quality of recruitment increases. It is likely that university educated workers were already recruited with care, so that there was not much room for efficient improvement there, which provides an explanation for the absence of a negative effect for the 0 to 11 months tenure group. To get a better understanding of the impact of the reform on the university/college educated workers, I fit the model to the Kaplan-Meier estimates of the firing hazard, as done previously for the full sample. Table 8 shows that after the reform, recruitment efforts remain roughly the same with about 63% of good matches. Note that this number is higher than for the whole sample before the reform (41%) and roughly equal to the sample mean after the reform. In other terms, university/college educated workers were indeed already recruited with much more care before the reform. After the reform, the recruitment effort for other employees catches up. The other salient finding of Table 8 is that employers have significantly increased monitoring efforts after the reform, with a standard error of the observation process going from 7.6 to 6.9. Lastly, allowing firing and hiring costs during the probationary period to change does not yield in this case a higher cost after the reform, but the cost seems to have slightly decreased. These findings altogether may explain why the WERS 2004 survey shows no evidence fore an increase in the use of tests for recruitment but does find an increase in the use of performance appraisals. Indeed, if tests and performance appraisals are mainly used for the more qualified workers, then these findings are consistent with the absence of change in recruitment efforts and increase in monitoring efforts found for the higher educated workers.

In conclusion, I do not find that males, older or more educated workers are most protected by the reform. Quite to the contrary, there is some evidence that females, younger and less educated workers are those who see the greatest reduction in their firing hazards. Moreover, heterogeneity in underlying parameters such as firing and hiring costs and the observability of performance does seem to be important, especially when considering different levels of education: thus, the reform has a different impact on the most educated workers when compared to other educational groups.

## 4.6 Impact on other separation hazards

To place firing in the context of other types of separation, I examine the hazard of any job separation after the reform (Figure 10). One can see that while all separations significantly decrease after the reform, they do not follow the same tenure pattern as firings, i.e. one does not see a trough in separations at around 24 months in the "before" period, and the hazards before and after become insignificantly different at tenure 30, and not tenure 24. Thus the shape of the *firing* hazard seems to be indeed determined by the existence of the right to claim unfair dismissal, while the overall separation hazard is not visibly affected by the consequences of that right. Moreover, to evaluate the global effect of the reform, it is interesting to note that while the firing hazard decreases, it is not the case that other types of separations increase at the same time so much as to imply no change in the overall separation hazard. In fact, the separation hazard is lower after the reform.

While the firing hazard has decreased after the reform, it is possible that firms have forced some workers to quit in order to avoid firing costs. These quits would then be disguised firings.

Figure 12 in Appendix 2 shows that the quit hazard did not increase after the reform. Because firms have increased their efforts towards higher match quality, one might expect to see a lower quit hazard, and so the fact that the latter only slightly diminishes may indicate that indeed some firms push the least productive workers to quit. Making the extreme assumption that all quits are in fact firings, I reproduce the analysis of Table 5. While I still find that the firing hazard has decreased, the decrease is now of lower magnitude, and it is only statistically significant for workers with 12 to 23 months of tenure (results not reproduced here). The assumption that all quits are disguised firings being extreme, I take the results of this analysis as showing that my findings are robust to shifts from firings to quits.

I next look at the impact of the reform on other key labor market outcomes such as training, wages or unemployment.

# 5 Impact on other labor market outcomes: training, wages, and unemployment duration

The analysis of the impact of the reform on other labor market outcomes is enlightening for two reasons. First, it allows for further investigation of the plausibility that firms have indeed increased their recruitment and monitoring efforts. Second, to better evaluate the overall welfare effect of the reform, one should look, beyond the effect on firing, at other positive or negative effects of the reform. In particular, it is essential to look at unemployment duration since theory predicts that with higher expected firing costs, one should see higher unemployment duration, and an increase in recruitment effort would only reinforce this effect.

However, the theory developed in section 3 does not directly generate predictions concerning the effects of the reform on labor market outcomes such as training, wages or unemployment duration. Indeed, that theory only applies to firing decisions taken by the firm. I will therefore have to use theoretical insights from other models of relevance in each particular case. However, because of the lack of appropriate theory and data, it is typically hard to find good control groups, and therefore estimates should be taken with caution.

## 5.1 Impact on wages and training

Theoretically, higher firing costs may increase or decrease wages. A first strand of theory argues that higher firing costs give a higher bargaining power to employed workers and so wages increase (Lindbeck and Snower, 2001). A second strand of theory argues that since workers value job security, they should accept lower wages (Summers, 1989). The relevant comparison in this case is workers with 12 to 23 months tenure versus workers with 24 to 48 months tenure. Indeed, workers with 12 to 23 months tenure are more expensive to fire after the reform, so this would imply an increase in their wages relative to the 24 to 48 months tenure group under the first theory and a decrease in wages under the second theory.

However, before I can test this effect on wages, I have to take into account the introduction of a National Minimum Wage, which came into force April 1st 1999. Studies of the effect of the minimum wage in the UK show that spillovers may have taken place on the wage distribution up to the first decile at most (Low Pay Commission, 2003). In order to eliminate the effects of the minimum wage, I look only at workers above the first decile of the wage distribution. Panel A of Table 9 shows the effect of the reform on wages of workers with different tenures: while if I use all workers, wages seem to have increased, and even significantly so for workers with 0 to 11 months tenure, when using only workers who were not affected by the minimum wage, this effect disappears. I therefore conclude that the reform had no significant effect on wages[20].

Training can be affected in two ways by a probationary period. First, higher firing costs after the probationary period can increase training in as much as it can be cheaper to train current marginal employees than to fire them and try to hire more productive employees. Thus, empirically firms who perceive higher firing costs are also more likely to train their workers (Pierre and Scarpetta, 2004). Another related theory is that firing costs increase implicit screening costs for all firms, which increases the value of the informational advantage of the current employer. Therefore the latter is more likely to provide training (Acemoglu and Pischke, 1998). These two theories

[20] The reader may wander at this point what happens to the main findings on the hazard of firing when restricting the sample to workers above the tenth decile of the wage distribution. Once one corrects for the sample selection this entails (in particular for the under-representation of high tenure workers among the observations where wage data is non-missing), the results are unaffected.

would imply in my setting that employees newly protected by the 1999 reform, that is employees with 12 to 23 months tenure, should receive more training after the reform. The training of employees with 0 to 11 months tenure may also increase because while the firing cost incurred by firms *if* they fire workers in that tenure range does not change after the reform, the *expected* firing cost does increase as the probationary period is now shorter. A second way in which firing costs could affect training is through the interaction between training and recruitment and monitoring efforts. First, training can select for a more productive and stable workforce: thus, Cappelli(2002) shows that employers who offer tuition assistance for their employees to go to college manage to select better quality employees who stay longer on the job. This would imply that training increases across the board after the reform as a strategy used by employers to generate better quality matches. Second, firms face a trade-off when deciding on the timing of training. On the one hand, training can be particularly beneficial at the beginning of the employment relationship because the worker can be better adapted to the job from the very beginning. On the other hand, the firm may not be willing to invest in workers whose quality is uncertain and whom it would be likely to fire later on. If, however, recruitment quality increases, training can take place earlier in the employment relationship. This predicts that workers with low tenure should receive more training. Third, training in the very beginning of the employment relationship may also be used as a screening and monitoring device, i.e. by training workers, firms may learn more about their ability than otherwise. From fitting the model in section 3, we know that employers have likely increased their monitoring efforts. This implies that training should have increased after the reform for the 0 to 11 months tenure group.

The proportion of workers who get training[21] has increased across the board after the reform (results not reproduced here), consistently with the idea that employers are trying to select for better matches, or that they train more

[21] Training is here « any training in the last four weeks ». Such training is paid for by the employer in a large majority of cases (71%). However, the information on who pays for training is only available for about a fourth of the sample, so I do not use it. The results are less significant but not different if I use only the sample where the information is available and I define training as "training paid for by the employer".

precisely because they manage to form better matches and the returns to training are increasing in match quality. Panel B of Table 9 documents the effect of the reform on firms' propensity to train their workers at different tenures. Workers with 12 to 23 months tenure do not get more training, while workers with 0 to 11 months tenure do. The impact of the reform on training is thus consistent with firms having increased their recruitment and monitoring efforts.

## 5.2 Impact on unemployment duration

There are three reasons why unemployment duration may increase after the reform. First, if expected firing costs increase, then labor demand may decrease, leading to higher unemployment duration. Second, even if labor demand does not decrease, firms' increased recruitment efforts could imply that it takes longer to pre-screen workers, and so unemployment duration should increase. Third, if match quality is not purely idiosyncratic but is also determined by general ability, and if moreover the current employer is better informed about the worker's general ability than the market, then a worker getting fired under higher firing costs sends a worse signal to the market. This would imply that workers fired between 1 and 2 years tenure after the reform should all other things equal have higher unemployment durations than workers fired between 1 and 2 years tenure before the reform[22].

Table 11 tries to identify the effects of the reform on unemployment duration. To perform this analysis, I use a sample of unemployed individuals in the sense of the International Labour Organization (ILO) from the same dataset I used for the employed. Summary statistics for this sample are provided in Table 10. To identify the effect of the reform on the duration of unemployment, I use two strategies. First, in panel A of Table 11, I look at the probability of finding a permanent job with more than 16 hours a week after the reform. The reform actually seems to have a positive effect on the probability of exiting unemployment towards a treated job. This is not to say that the reform has actually increased this probability by 11%, but it seems that at least any negative impact of the reform has been overpowered by

[22] Unfortunately, for lack of a long enough follow up period, it is not possible to properly test this specific hypothesis.

otherwise positive trends. Moreover, increased match quality may create positive externalities so that the supply of treated jobs may increase despite the cost to individual firms of increasing match quality. Consistent with this finding, I observe that the proportion of permanent jobs among jobs with 16 or more hours a week steadily increases, just as much before as after the June 1999 reform (Figure 11). Therefore, it does not seem that the reform incited employers to substitute away from full-time permanent jobs.

A second strategy I use in panel B of Table 11 is to look at the exit towards any job and use the difference between those looking for full-time jobs and the others. Because the unfair dismissal provisions only apply to full-time jobs, we expect that workers looking for full-time jobs take longer to find a job relative to other unemployed workers. Note that part-time workers are actually a good control group because since the Part-time Workers (Prevention of Less Favourable Treatment) Regulations 2000, which came into force on July 1st 2000, they have the same rights as full-time workers in most areas, except precisely for this right to claim unfair dismissal.

The table shows no negative effect of the reform on the duration of unemployment for workers looking for full-time jobs relative to the others[23]. Quite to the contrary, the reform seems to have a positive effect implying that workers looking for a full-time job are 9% more likely to exit unemployment after the reform. Once again, this is not to say that the reform had a causal effect, but that any negative effects have been overrun by stronger positive effects. In conclusion, the reform has no discernable negative effect on the duration of unemployment or on the relative supply of permanent jobs with more than 16 hours a week, implying that any negative effects have been overpowered by positive ones.

[23] Note that being part-time or full-time is left by the LFS to the subjective appreciation of the worker. In practice, 37.55% of workers who say they work part-time and are in permanent jobs work 15 hours or less, and 45.45% work 16 or less hours. Therefore, some of the "part-timers" are de facto also affected by the unfair dismissal provision. This means that any negative effect of the reform will be underestimated if one compares workers looking for full-time jobs versus the others.

# 6 Conclusion and possible extensions

Using a learning model, I have shown that the existence of a probationary period influences firms' firing pattern so that, all else equal, there is a peak in the firing hazard just before the end of the probationary period and a trough right after the end of the probationary period. This effect is smaller the smaller the difference between firing costs before and after the end of the probationary period. The empirical analysis showed that shortening the qualifying period for the right to claim unfair dismissal reduced the hazard of firing for newly covered workers, but also for workers with lower tenure, reflecting firms' increase in recruitment efforts. Firms have also increased their monitoring efforts and their investment in training after the reform.

These results are only partially consistent with the predictions of the British labor government about the impact of the reform (section 2). First, they predicted that it would encourage workers to change jobs, leading to a more flexible labor market. This is not the case however as quits and overall separations have actually decreased. Second, they predicted that employers would adopt better employment practices, thus increasing productivity: this seems to have happened since employers are more careful about whom they hire, they monitor their workers better, and they train them more. Lastly, the government thought that one year is enough time for the initial screening of workers: this does not seem to be confirmed by the data, since the reform prompted firms to change their human resource management policies, precisely to limit the need for firing past one year of tenure.

These results on the British reform are of particular interest for the evaluation of the new CNE contract in France and the longer probationary period proposed in Germany by the Merkel government: thus, lengthening the probationary period should increase firing, decrease match quality, and have a limited impact on employment. However, these predictions are based on direct extrapolation of the British results, and do not take into account the specificities of the French or German economies. I therefore plan to evaluate these reforms directly as data becomes available.

In the debate about the effects of firing costs, this work has shown that the British reduction in the probationary period, and the associated increase in expected firing costs, did not have any discernable effect on employment or the duration of unemployment, while it likely increased productivity via better matches and more training.

This paper could be extended along several lines. First, to better understand the mechanisms at play, it would be helpful to examine countries with different lengths of the probationary period and different firing costs. The United Kingdom is indeed a special case: while its employment law is very similar in structure to that of the countries from continental Europe, firing costs are much lower on average. Examining more typical European countries such as France or Germany should thus shed more light on how a probationary period affects firms' behavior and labor market outcomes in the European institutional context. Second, the model used here could be applied to other questions. For example, I have shown how the distributions of firms' beliefs about workers' productivity evolve with tenure: this may have important implications for the wage distribution by tenure.

In general, it would be useful to further investigate how the widely spread institution of a probationary period can solve the trade-offs policy makers face when deciding on firing costs. While I have shown some ways in which a probationary period can affect economic efficiency, i.e. for example by influencing firms' investments in match quality and human capital, this paper sheds little light on how this institution affects labor demand or interacts with the business cycle. The analysis of general equilibrium and business cycle effects of tenure-dependant job protection is thus a promising avenue for future research.

**Bibliography**


ACEMOGLU D., PISCHKE J. S. (1998), "Why Do Firms Train? Theory and Evidence", *Quarterly Journal of Economics*, volume 113, pp. 79-119.

AUTOR D. H., DONOHUE III J. J., SCHWAB S. J. (2004), "The costs of wrongful discharge laws", September 2004.

BAKER M., REA S. A. JR. (1998), "Employment Spells and Unemployment Insurance Eligibility Requirements", *The Review of Economics and Statistics*, Vol. 80, No. 1 (Feb., 1998), 80-94.

BERTOLA G. (1992), "Labor Turnover Costs and Average Labor Demand," *Journal of Labor Economics*, 10:4 (October 1992), pp.389-411.

BLACKBURN, R., HART M. (2002), "Small firms' awareness and knowledge of individual employment rights", *Department of Trade and Industry*, DTI Employment Relations Research Series No. 14, http://www.dti.gov.uk/er/emar/errs14.pdf.

BLANCHARD, O., KATZ L.F. (1997), "What We Know and Don't Know About the Natural Rate of Unemployment", *Journal of Economic Perspectives*, vol. 11, no. 1, pp. 51-72.

BLANCHARD, O., LANDIER, A. (2002), "The Perverse Effect of Partial Labour Market Reform: Fixed-Term Contracts in France", *Economic Journal*, vol. 112, no. 480, pp. F214-244.

CAPPELLI P. (2002), "Why Do Employers Pay for College?", NBER Working Paper No. 9225.

CHRISTOFIDES L.N., MCKENNA C.J. (1996), "Unemployment Insurance and Job Duration in Canada", *Journal of Labour Economics*, Vol. 14, No. 2 (Apr., 1996), 286-312.

DAVIES P.L., FREEDLAND M. (1993), *Labour legislation and public policy: a contemporary history*. Oxford: Clarendon Press, 1993.

DJANKOV S., LA PORTA R., LOPEZ-DE-SILANE F., SHLEIFER A., BOTERO J. (2004), "The Regulation of Labour", *Quarterly Journal of Economics*, Vol. 119, no. 4, pp. 1339-1382.

FARBER H. S. (1994), "The Analysis of Interfirm Worker Mobility", *Journal of Labor Economics*, Vol 12., 554-593.

GIBBONS, R, KATZ, L.F. (1991), "Layoffs and Lemons", *Journal of Labor Economics*, Vol. 9, No. 4 (October, 1991), 351-80.

GIBBONS, R, KATZ, L.F., LEMIEUX, T., PARENT D. (2005), "Comparative Advantage, Learning, and Sectoral Wage Determination", *Journal of Labor Economics*, Vol. 23, No. 4 (October, 2005), forthcoming.

HECKMAN J., PAGÉS C. (2003), "Law and Employment: Lessons from Latin America and the Caribbean", NBER Working Paper 10129, December 2003

HUNT J. (2000), "Firing Costs, Employment Fluctuations and Average Employment: An Examination of Germany", *Economica* 2000, 67, 177-202.

JOVANOVIC B. (1979), "Job Matching and the Theory of Turnover", *The Journal of Political Economy*, Vol 87., 972-990.

KERSLEY B., ALPIN C., FORTH J., BRYSON A., BEWLEY H., DIX G., OXENBRIDGE S. (2005) "Inside the Workplace: First Findings from the 2004 Workplace Employment Relations Survey (WERS 2004)", DTI, http://www.dti.gov.uk/er/insideWPfinalwebJune.pdf.

KUGLER A. (1999), "The Impact of Firing Costs on Turnover and Unemployment: Evidence from the Colombian Labor Market Reform", *International Tax and Public Finance*, August 1999, 6(3): 389-410.

KUGLER A., PICA G., "Effects of Employment Protection on Job and Worker Flows: Evidence from the 1990 Italian Reform" (with Giovanni Pica), Mimeo, May 2005.

LANCASTER, T. (1990), *The Econometric Analysis of Transition Data*, Econometric Society Monographs no. 17. Cambridge: Cambridge University Press, 1990.

LAZEAR E. (1990), "Job Security Provisions and Employment" *The Quarterly Journal of Economics,* Vol. 105, No. 3 (August), pp. 699-726 .

LINDBECK, A., SNOWER D. J. (1990), "Insiders versus outsiders", *The Journal of Economic Perspectives*, Vol. 15, No. 1 (Winter, 2001), pp. 165-188.

LOW PAY COMMISSION (2003), "The National Minimum Wage: Fourth Report of the Low Pay Commission, Building on Success", http://www.lowpay.gov.uk/lowpay/report/pdf/lowpay-nmw.pdf

MALCOMSON J. (1999), "Individual Employment Contracts," The Handbook of Labor Economics, Orley Ashenfelter and David Card (eds.) Vol 3, Elsevier Science, 1999.

MARINESCU I. (2006B), "The determinants of the separation hazard in a model with learning and time-varying match quality," working paper.

MOFFITT, R. JOVANOVIC, B. (1990), "An Estimate of a Sectoral Model of Labor Mobility", *The Journal of Political Economy*, Vol. 98, No. 4 (Aug., 1990), pp. 827-852

NAGYPAL E. (2004), "Learning-by-Doing Versus Learning About Match Quality: Can We Tell Them Apart?", working paper, September 2004.

NICKELL S., NUNZIATA L., OCHEL W. (2005), "Unemployment in the OECD since the 1960's. What do we know?," *Economic Journal*, 115 (January), pp. 1-27 .

OECD (1999), "Employment Protection and Labour Market Performance", Employment Outlook, 1999.

PIERRE G., SCARPETTA S. (2004), "Employment Regulations through the Eyes of Employers: Do they Matter and How Do Firms Respond to Them?," Policy Research Working Paper Series 3463, The World Bank.

SCHOENBERG U. (2004), "Testing for Asymmetric Employer Learning", Working Paper, June 2004.

SUMMERS L. H. (1989), "Some Simple Economics of Mandated Benefits", *The American Economic Review*, Vol. 79, No. 2, Papers and Proceedings of the Hundred and First Annual Meeting of the American Economic Association (May, 1989), pp. 177-183

Table 1 : parameters used to compute results in the benchmark case

| **Parameters** | **Values** |
|---|---|
| Discount factor $\delta$ | .995 |
| Initial proportion of good matches $q$ | .5 |
| Standard error of signal $\sigma$ | 4 |
| Firing costs $c$ | 7 |
| Maximal tenure | 200 |

Figure 1 : the effect of a probationary period on the firing hazard

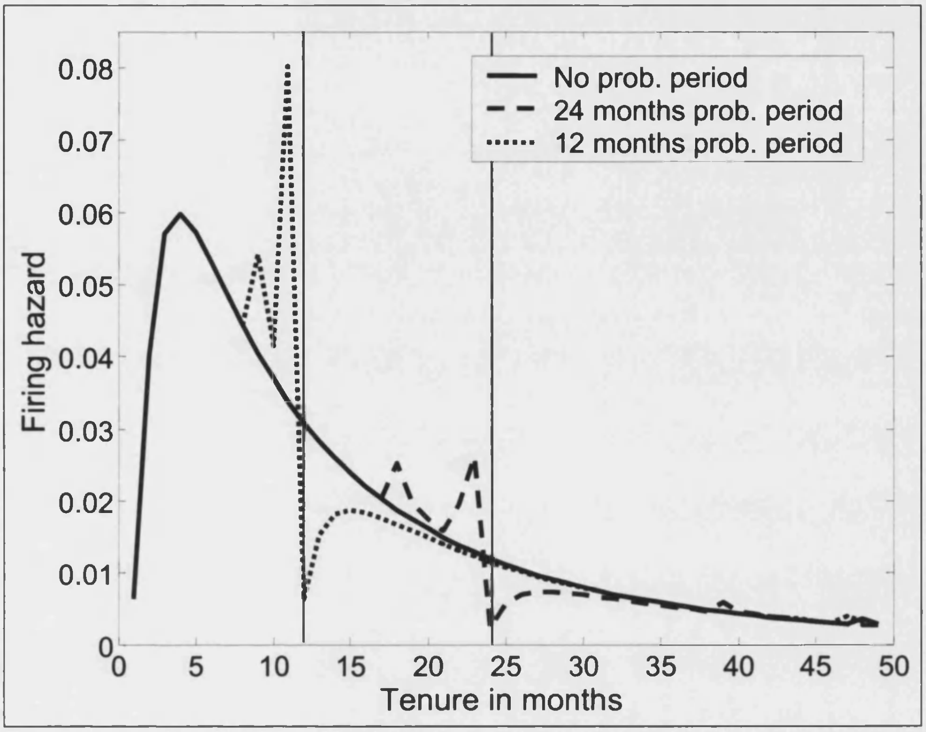


Figure 2: belief distributions in the reference case

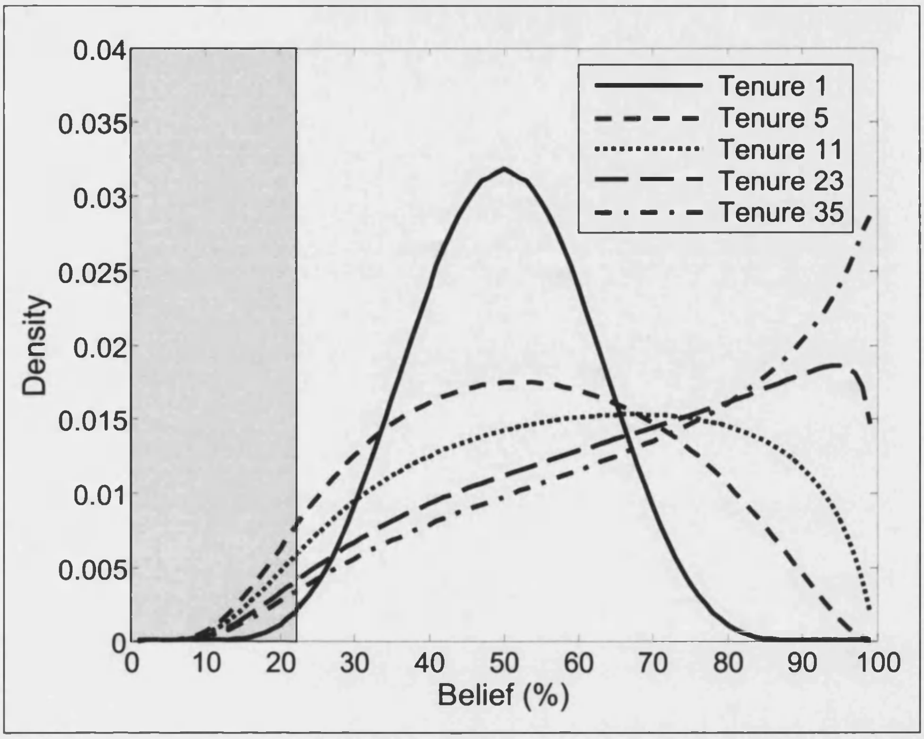


Figure 3: the effect of an increase in recruitment effort or monitoring intensity after the reform

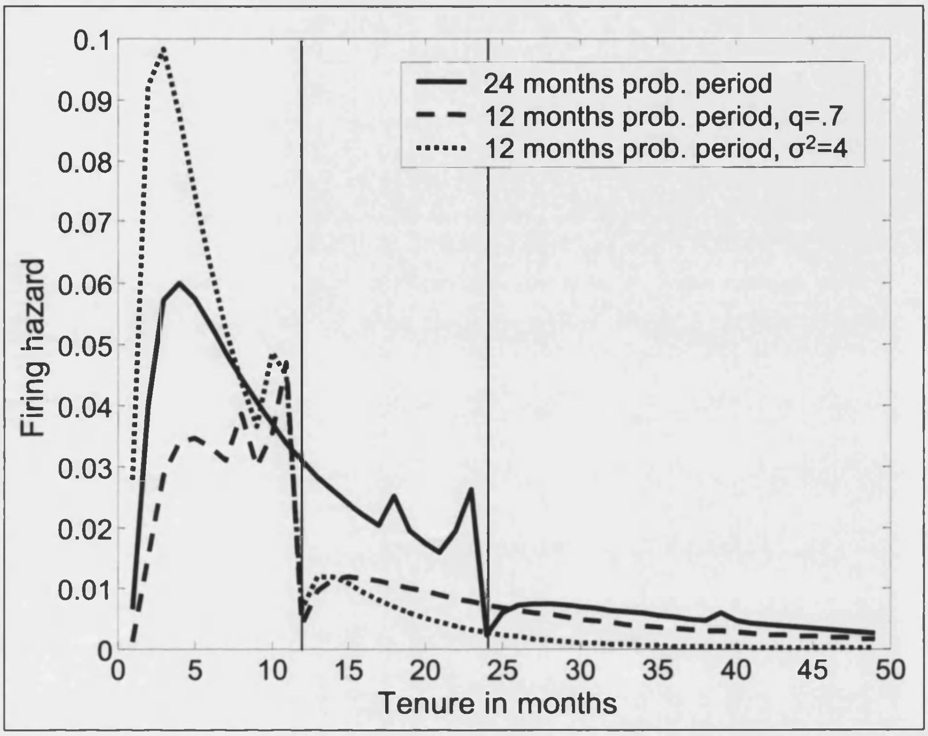


Figure 4: belief distributions in the case of an increase in recruitment effort after the reform

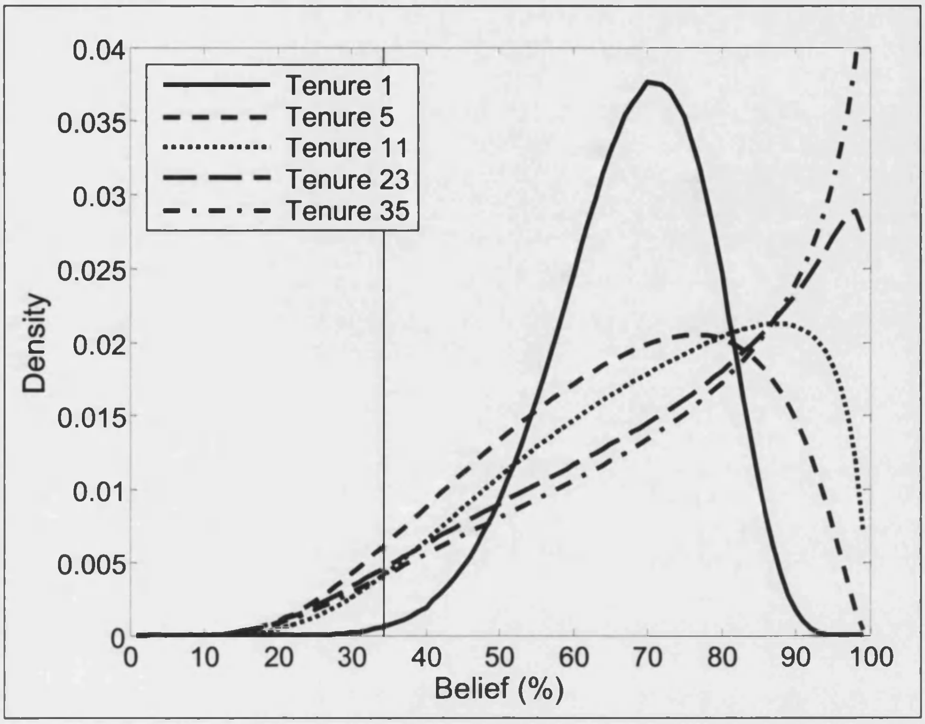


Figure 5: distributions of belief in the case of an increase in monitoring intensity after the reform

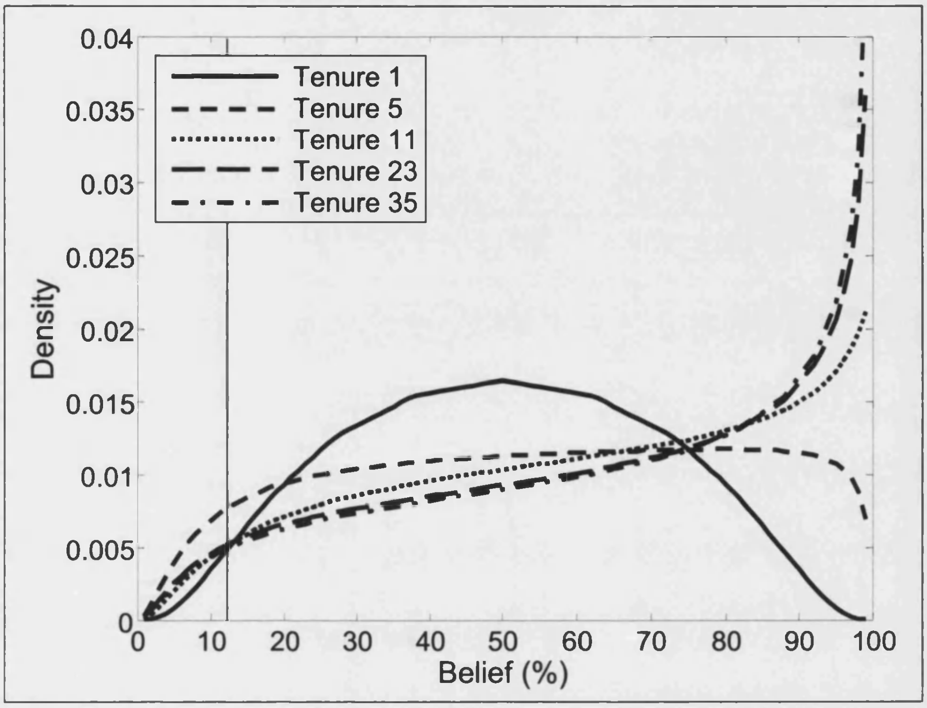
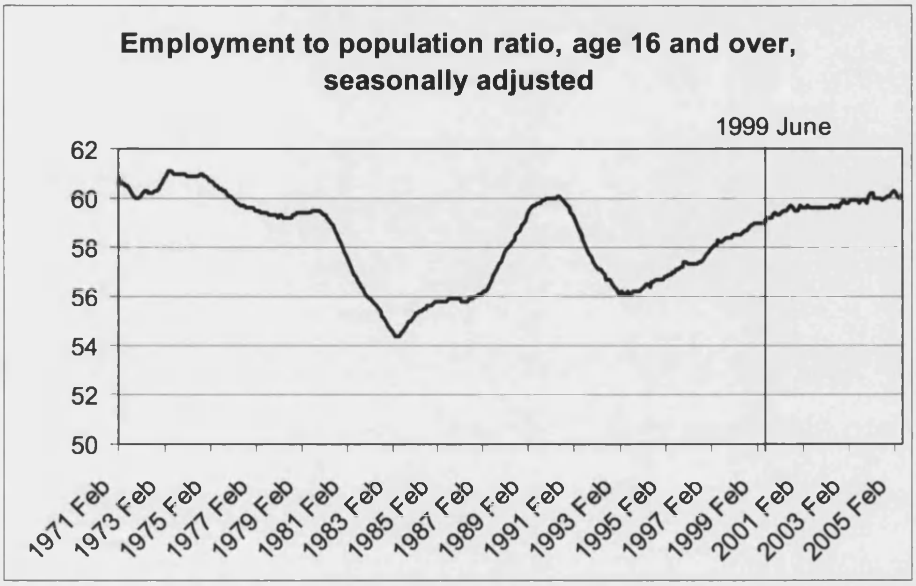


Table 2: Effects predicted by the model for a reduction in the probation period from 24 to 12 months

| | No change in recruitment effort or monitoring | Increase in recruitment effort | Increase in monitoring |
|---|---|---|---|
| **0 to a few months tenure** | NONE | --- | +++ |
| **12 to 24 months tenure** | --- | --- | --- |
| **24 months tenure and more** | ~NONE | --- | --- |

Figure 6: The evolution of the employment to population ratio



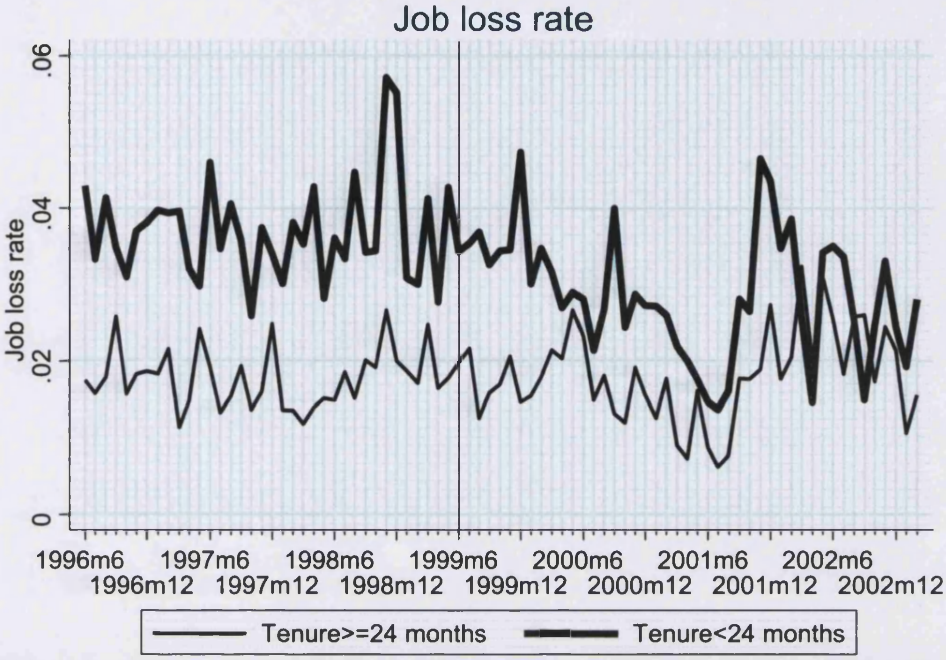

Source: UK National Statistics, MGSR series, computed from the Labour Force Survey.

Table 3: Summary statistics for the sample of permanent full-time employees

| | Obs. | Mean | Std. | Min | Max |
|---|---|---|---|---|---|
| **Macro situation** | | | | | |
| **Unemployment rate (claimant count)** | 436867 | 3.954 | 1.706 | 1.5 | 11.7 |
| **Reason for leaving last job** | | | | | |
| **dismissed** | 39954 | 0.030 | 0.172 | 0 | 1 |
| **made redundant,voluntary redundancy** | 39954 | 0.183 | 0.389 | 0 | 1 |
| **temporary job ended** | 39954 | 0.034 | 0.180 | 0 | 1 |
| **resigned** | 39954 | 0.358 | 0.479 | 0 | 1 |
| **gave up work for health reasons** | 39954 | 0.046 | 0.209 | 0 | 1 |
| **took early retirement** | 39954 | 0.024 | 0.153 | 0 | 1 |
| **retired** | 39954 | 0.026 | 0.160 | 0 | 1 |
| **family, personal reason** | 39954 | 0.074 | 0.261 | 0 | 1 |
| **left for some other reason** | 39954 | 0.225 | 0.417 | 0 | 1 |
| **Job characteristics** | | | | | |
| **Tenure** | 436097 | 98.456 | 101.866 | 0 | 652 |
| **Usual hours worked per week** | 433442 | 36.596 | 8.948 | 16 | 97 |
| **Gross weekly wage in pounds** | 167695 | 333.354 | 282.744 | 1 | 44000 |
| **Log real hourly wage** | 166926 | -2.633 | 0.571 | -8.792 | 2.342 |
| **Job training** | 435358 | 0.287 | 0.452 | 0 | 1 |
| **Person characteristics** | | | | | |
| **Female** | 436867 | 0.460 | 0.498 | 0 | 1 |
| **Married and cohabiting** | 436867 | 0.580 | 0.494 | 0 | 1 |
| **Age** | 436867 | 38.850 | 11.566 | 16 | 64 |
| **Less than high school educated** | 436771 | 0.247 | 0.432 | 0 | 1 |
| **University educated** | 436771 | 0.278 | 0.448 | 0 | 1 |
| **Occupation categories** | | | | | |
| **Manager** | 436690 | 0.161 | 0.368 | 0 | 1 |
| **Professional** | 436690 | 0.111 | 0.314 | 0 | 1 |
| **Associate professional and technical** | 436690 | 0.121 | 0.326 | 0 | 1 |
| **Administrative and secretarial** | 436690 | 0.159 | 0.366 | 0 | 1 |
| **Skilled trades occupations** | 436690 | 0.107 | 0.309 | 0 | 1 |
| **Personal service occupations** | 436690 | 0.090 | 0.285 | 0 | 1 |
| **Sales and customer service occupations** | 436690 | 0.073 | 0.260 | 0 | 1 |
| **Process, plant and machine operatives** | 436690 | 0.098 | 0.297 | 0 | 1 |
| **Elementary occupations** | 436690 | 0.081 | 0.273 | 0 | 1 |
| **Employer characteristics** | | | | | |
| **Private sector employer** | 435832 | 0.643 | 0.479 | 0 | 1 |
| **Manufacturing or construction sector** | 436699 | 0.238 | 0.426 | 0 | 1 |
| **Administration sector** | 436699 | 0.044 | 0.205 | 0 | 1 |

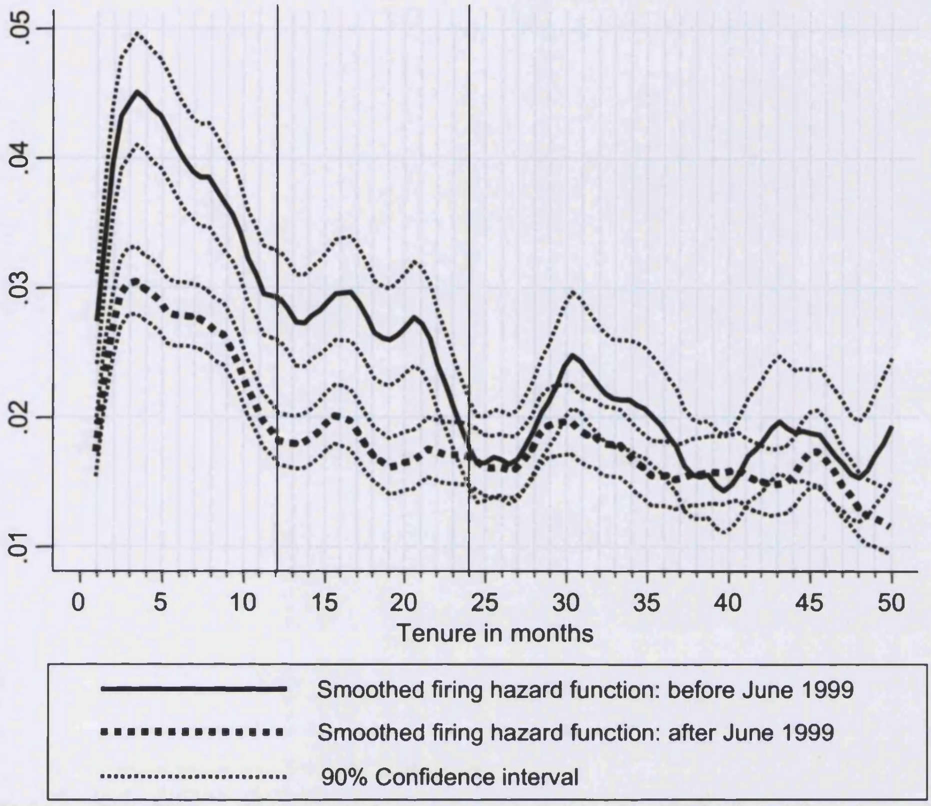
Notes: The sample is restricted to persons who are employed in the first quarter, in a permanent job, and usually working 16 or more hours a week. Only the first observation for each job (as defined by the hiring date) is kept.
Source: Labour Force Survey Two-Quarter Longitudinal Dataset (www.data-archive.ac.uk). For the unemployment rate, UK National Statistics, Time Series data [NS TSD], Regional claimant count rate, non seasonally adjusted, series code EGU4.

Figure 7: Job loss rate




Notes: the job loss rate is calculated as the number of workers who were dismissed or made redundant between the first and the second interview quarter over the total

Figure 8: Kaplan-Meier estimates of the firing hazard before and after the reform




Notes: The figure plots smoothed non-parametric Kaplan-Meier firing hazard estimates. Firing is defined as dismissing or making redundant a worker. The sample is restricted to persons who are employed in the first quarter, in a permanent job, and usually working 16 or more hours a week. Only the first observation for each person is kept.

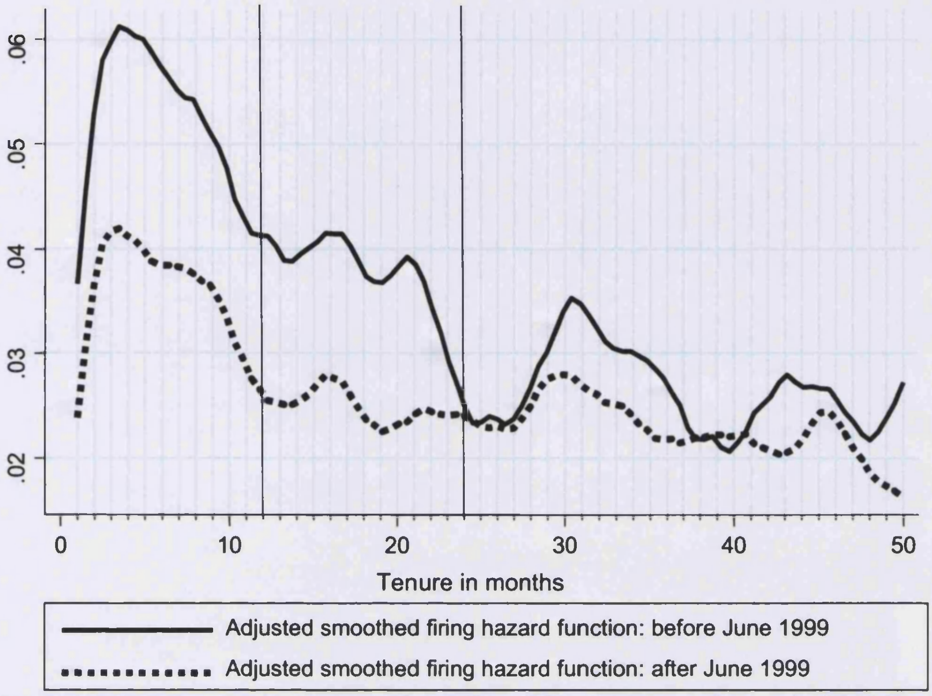
Source: Labour Force Survey Two-Quarter Longitudinal Dataset (www.data-archive.ac.uk).

Table 4: parameters of the calibrated model

| | Before June 1999 | After June 1999 | After June 1999 |
|---|---|---|---|
| Length of probationary period | 24 months | 12 months | 12 months |
| q | **0.414** | **0.630** | **0.624** |
| σ | **5.706** | **5.554** | **5.567** |
| c0 | **6.602** | 6.602 | **6.782** |
| c1 | **6.800** | 6.800 | 6.800 |
| Discount factor | 0.995 | 0.995 | 0.995 |
| Maximal tenure | 200 months | 200 months | 200 months |

Notes: The bold numbers are those that were calibrated, while the other numbers were taken as parameters. c0 is the firing cost during the probationary period and c1 is the firing cost after the probationary period. The model is calibrated to best fit the 36 first months of the empirical hazard function in the case of a 24 months probationary period, and the 24 first months of the empirical hazard function in the case of a 12 months probationary period.

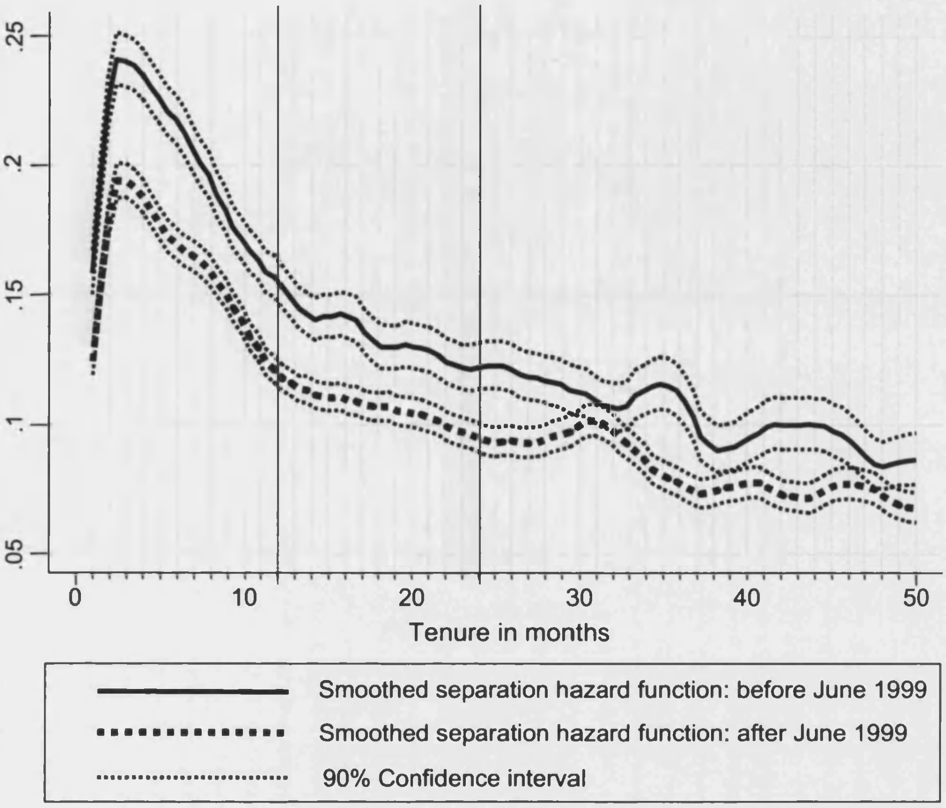
Figure 9: Adjusted estimates of the firing hazard before and after the reform




Notes: The plots are the smoothed baseline firing hazards after a stratified Cox regression model where "before" and "after" are the two strata. Firing is defined as dismissing or making redundant a worker. The control variables included in the Cox regression are: the regional unemployment rate in the month under consideration, age, gender, education, occupation, sector (public or private), industry. The graph is then plotted at the median values of these variables (when the latter are categorical and cover more than one category, the most frequent category is used). The sample is restricted to persons who are employed in the first quarter, in a permanent job, and usually working 16 or more hours a week. Only the first observation for each person is kept.

Source: Labour Force Survey Two-Quarter Longitudinal Dataset (www.data-archive.ac.uk). For the unemployment rate: UK National Statistics, Time Series data [NS TSD], Regional claimant count rate, non seasonally adjusted, series code EGU4.

Figure 10: Kaplan-Meier estimates of the separation hazard before and after the reform




Notes: same as Figure 8, except that the failure event here is any job separation, instead of dismissals or redundancies only.

Table 5: Impact of the reform on the hazard of firing by tenure

| | Post reform period: After June 1999 | | Post reform period: After June 1999, excluding May 1998 to May 2000 | |
|---|---|---|---|---|
| | Control group: 24 months and more tenure | Control group: 24-48 months tenure | Control group: 24 months and more tenure | Control group: 24-48 months tenure |
| | A. Basic tenure categories | | | |
| **0 to 11 months tenure** | -0.182 | -0.267 | -0.158 | -0.282 |
| | (0.050)*** | (0.072)*** | (0.060)*** | (0.087)*** |
| **12 to 23 months tenure** | -0.205 | -0.290 | -0.202 | -0.326 |
| | (0.065)*** | (0.083)*** | (0.078)** | (0.100)*** |
| | B. Detailed tenure categories | | | |
| | Control group: 26 months and more tenure | Control group: 26-48 months tenure | Control group: 26 months and more tenure | Control group: 26-48 months tenure |
| **0 to 4 months tenure** | -0.135 | -0.223 | -0.071 | -0.195 |
| | (0.069)* | (0.088)** | (0.085) | (0.108)* |
| **5 to 11 months tenure** | -0.203 | -0.292 | -0.190 | -0.313 |
| | (0.060)*** | (0.082)*** | (0.073)*** | (0.099)*** |
| **12 to 17 months tenure** | -0.203 | -0.292 | -0.188 | -0.311 |
| | (0.081)** | (0.098)*** | (0.099)* | (0.120)*** |
| **18 to 21 months tenure** | [illegible] | [illegible] | [illegible] | [illegible] |
| | (0.231) | (0.240) | (0.194) | (0.201) |
| **Number of observations** | 335782 | 335782 | 430604 | 430604 |

Notes: The coefficients reported are the interactions between tenure categories and "after". Cox proportional hazard models are used.

Table 6: Impact of the reform on the firing hazard by gender and age

| | A. Gender | |
|---|---|---|
| | Males | Females |
| **0 to 11 months tenure** | -0.223 | -0.392 |
| | (0.106)** | (0.153)** |
| **12 to 23 months tenure** | -0.307 | -0.348 |
| | (0.122)** | (0.177)** |
| **Number of observations** | 180899 | 154883 |
| | B. Age | |
| | Age<40 | Age>=40 |
| **0 to 11 months tenure** | -0.299 | -0.263 |
| | (0.110)*** | (0.146)* |
| **12 to 23 months tenure** | -0.421 | -0.150 |
| | (0.127)*** | (0.167) |
| **Number of observations** | 174762 | 204265 |

Notes: The coefficients reported are the interactions between tenure categories and the "after June 1999, excluding May 1998 to May 2000" dummy. Cox proportional hazard models are used.
Robust standard errors clustered by person in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.
The control group is 24 to 48 months tenure.
All regressions include the following controls: tenure categories dummies (same as listed in the table), unemployment rate, married and cohabiting dummy, age, 2 education dummies, 8 occupational dummies, private sector dummy, manufacturing and construction dummy, administration dummy, 3 quarters dummies, year dummies (years are June to May). Regressions in panel B also include a female dummy.
The sample is restricted to persons who are employed in the first quarter, in a permanent job, and usually working 16 or more hours a week. Only the first observation for each job (as defined by the hiring date) is kept.
Source: Labour Force Survey Two-Quarter Longitudinal Dataset (www.data-archive.ac.uk). For the unemployment rate: UK National Statistics, Time Series data [NS TSD], Regional claimant count rate, non seasonally adjusted, series code EGU4.

Table 7: Impact of the reform on the firing hazard by education level

| | Less than high school | High school but less than college | University/College educated |
|---|---|---|---|
| **0 to 11 months tenure** | -0.358 | -0.319 | 0.145 |
| | (0.153)** | (0.125)** | (0.215) |
| **12 to 23 months tenure** | -0.297 | -0.429 | -0.166 |
| | (0.177)* | (0.148)*** | (0.220) |
| **Number of observations** | 81712 | 159224 | 94846 |

Notes: The coefficients reported are the interactions between tenure categories and the "after June 1999, excluding May 1998 to May 2000" dummy. Cox proportional hazard models are used.

Robust standard errors clustered by person in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

The control group is 24 to 48 months tenure.

All regressions include the following controls: tenure categories dummies (same as listed in the table), unemployment rate, female dummy, married and cohabiting dummy, age, 8 occupational dummies, private sector dummy, manufacturing and construction dummy, administration dummy, 3 quarters dummies, year dummies (years are June to May).

The sample is restricted to persons who are employed in the first quarter, in a permanent job, and usually working 16 or more hours a week. Only the first

y the hiring date) is kept.

rce: Labour Force Survey Two-Quarter Longitudinal Dataset (www.data-
hive.ac.uk). For the unemployment rate: UK National Statistics, Time Series data
TSD], Regional claimant count rate, non seasonally adjusted, series code EGU4.

So
arc
[NS

Table 8: Parameters of the calibrated model for college/university educated workers

| | Before June 1999 | After June 1999 | After June 1999 |
|---|---|---|---|
| Length of probationary period | 24 months | 12 months | 12 months |
| q | **0.631** | **0.633** | **0.633** |
| σ | **7.616** | **6.901** | **6.902** |
| c0 | **7.778** | 7.778 | **7.761** |
| c1 | **7.801** | 7.801 | 7.801 |
| Discount factor | 0.995 | 0.995 | 0.995 |
| Maximal tenure | 200 months | 200 months | 200 months |

Notes: The bold numbers are those that were calibrated, while the other numbers were taken as parameters. c0 is the firing cost during the probationary period and c1 is the firing cost after the probationary period. The model is calibrated to best fit the 36 first months of the empirical hazard function in the case of a 24 months probationary period, and the 24 first months of the empirical hazard function in the case of a 12 months probationary period.

Table 9: Impact of the reform on wages and training

| | A.Log real hourly wage | | B. Training |
|---|---|---|---|
| | All workers | Workers above the 1st decile | |
| **0 to 11 months tenure** | 0.016 | -0.002 | 0.017 |
| | (0.009)* | (0.008) | (0.007)** |
| **12 to 23 months tenure** | 0.012 | 0.009 | -0.008 |
| | (0.008) | (0.008) | (0.007) |
| **R-squared** | 0.45 | 0.44 | |
| **Number of observations** | 126855 | 106934 | 338238 |

Notes: The coefficients reported are the interactions between tenure categories and the "after June 1999, excluding May 1998 to May 2000" dummy. Panels A reports results from OLS regressions. Panel B reports the marginal effects from a probit model; while the marginal interactions effects are not properly calculated by the dprobit Stata command, the coefficients from a linear probability model are quasi identical.
Robust standard errors clustered by person in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.
The control group is 24 to 48 months tenure.
All regressions include the following controls: tenure categories dummies (same as [redacted]

Table 10: Summary statistics for the sample of ILO unemployed

| | Obs. | Mean | Std. | Min | Max |
|---|---|---|---|---|---|
| **Macro situation** | | | | | |
| **Unemployment rate (claimant count)** | 38004 | 4.437 | 1.827 | 1.5 | 11.7 |
| **Unemployment spell characteristics** | | | | | |
| **Unemployment duration** | 38004 | 31.775 | 52.734 | 0 | 482 |
| **Seeking full-time employee job** | 38004 | 0.513 | 0.500 | 0 | 1 |
| **Person characteristics** | | | | | |
| **Female dummy** | 38004 | 0.408 | 0.491 | 0 | 1 |
| **Married and cohabiting dummy** | 38004 | 0.372 | 0.483 | 0 | 1 |
| **Age** | 38004 | 36.076 | 12.667 | 15 | 64 |
| **Less than high school educated dummy** | 37997 | 0.395 | 0.489 | 0 | 1 |
| **University educated dummy** | 37997 | 0.156 | 0.363 | 0 | 1 |

Notes: The sample is restricted to persons who are ILO unemployed in the first quarter and whose date of leaving their previous job is known. Only the first observation for each unemployment spell (as defined by the date when the last job was left) is kept.
Source: Labour Force Survey Two-Quarter Longitudinal Dataset (www.data-archive.ac.uk). For the unemployment rate: UK National Statistics, Time Series data [NS TSD], Regional claimant count rate, non seasonally adjusted, series code EGU4.

Table 11: Impact of the reform on the duration of unemployment

| | A. Exit unemployment towards a permanent job with more than 16 hours a week | B. Exit unemployment towards any job |
|---|---|---|
| **After** | 0.107<br>(0.031)*** | |
| **Looking preferably for full-time employee job** | | 0.005<br>(0.025) |
| **Looking preferably for full-time employee job*After** | | 0.086<br>(0.030)*** |
| **Number of observations** | 27966 | 27956 |

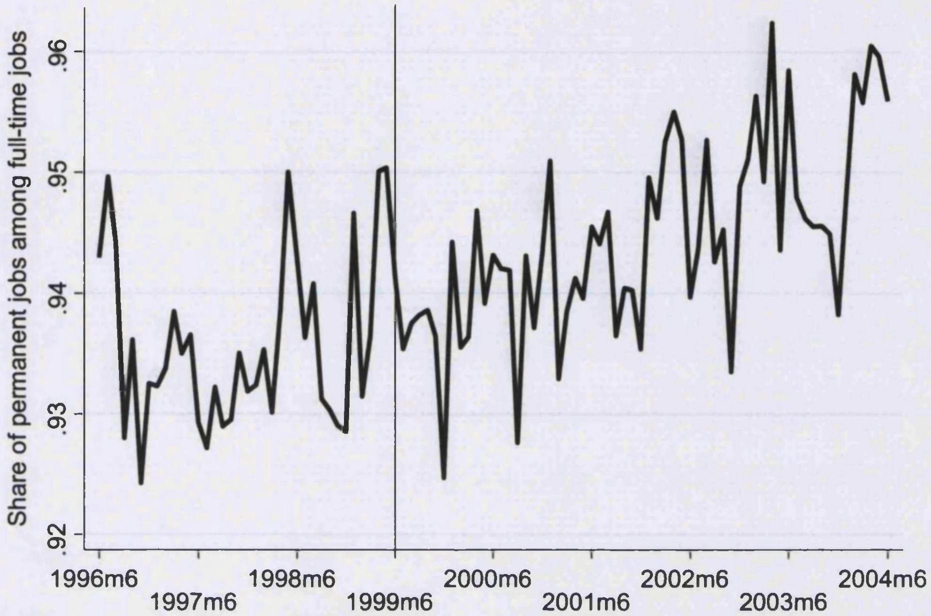
Notes: Cox proportional hazard models are used. After is defined as "after June 1999, excluding May 1998 to May 2000".

Robust standard errors clustered by person in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

All regressions include the following controls: unemployment rate, female dummy, married and cohabiting dummy, age, 2 education dummies, 3 quarters dummies. Regressions in panel B also include dummies for types of job looked for, dummies for types of job looked for interacted with "after", and year dummies (years are June to May).

The sample is restricted to persons who are ILO unemployed in the first quarter and

Figure 11: Evolution of the proportion of permanent jobs among full-time jobs




Notes: The sample is restricted to persons who are employed in the first quarter and usually working 16 or more hours a week. Only the first observation for each job (as defined by the hiring date) is kept.

Source: Labour Force Survey Two-Quarter Longitudinal Dataset (www.data-archive.ac.uk).

APPENDIX 1 : Equations for the firm's belief about match quality, and the firing hazard

**Belief**

The sum of observations out of $t$ periods is described, under my hypotheses, by a normal distribution. Let $g_g(s,t)$ be the probability of getting a sum $s$ of observations at tenure $t$ when the true match quality is good: the distribution is normal with mean $t$ and variance $t.\sigma^2$. Symmetrically $g_b(s,t)$ is normal with mean $-t$ and variance $t.\sigma^2$. Using Bayes' rule we can then compute all possible beliefs. We have:

$$b(s,t) = \frac{q.g_g(s,t)}{q.g_g(s,t) + (1-q).g_b(s,t)}$$

It turns out that $t$ drops out and the formula simplifies to:

$$b(s,t) = \frac{q.\exp\left(\frac{s}{\sigma^2}\right)}{q.\exp\left(\frac{s}{\sigma^2}\right) + (1-q).\exp\left(-\frac{s}{\sigma^2}\right)}$$

**Firing hazard**

Let $f_t(s)$ be the density of matches with sum of observations $s$ at time $t$.

The initial values are:

$f_0(0) = 1$

$\forall s \neq 0, f_0(s) = 0$

Let $p(s \mid s_1)$ be the probability density of getting a total sum of observations $s$ when at the previous period the total sum of observations was $s_1$.

$$p(s \mid s_1) = \frac{b(s_1)}{\sigma\sqrt{2\pi}}\exp\left(\frac{-(s-s_1-b(s_1))^2}{2\sigma^2}\right) + \frac{1-b(s_1)}{\sigma\sqrt{2\pi}}\exp\left(\frac{-(s-s_1+1-b(s_1))^2}{2\sigma^2}\right)$$

The evolution of the density of matches is given by the following recursion equation, where $s(\tau(t))$ is the sum of observations corresponding to the belief threshold at tenure $t$:

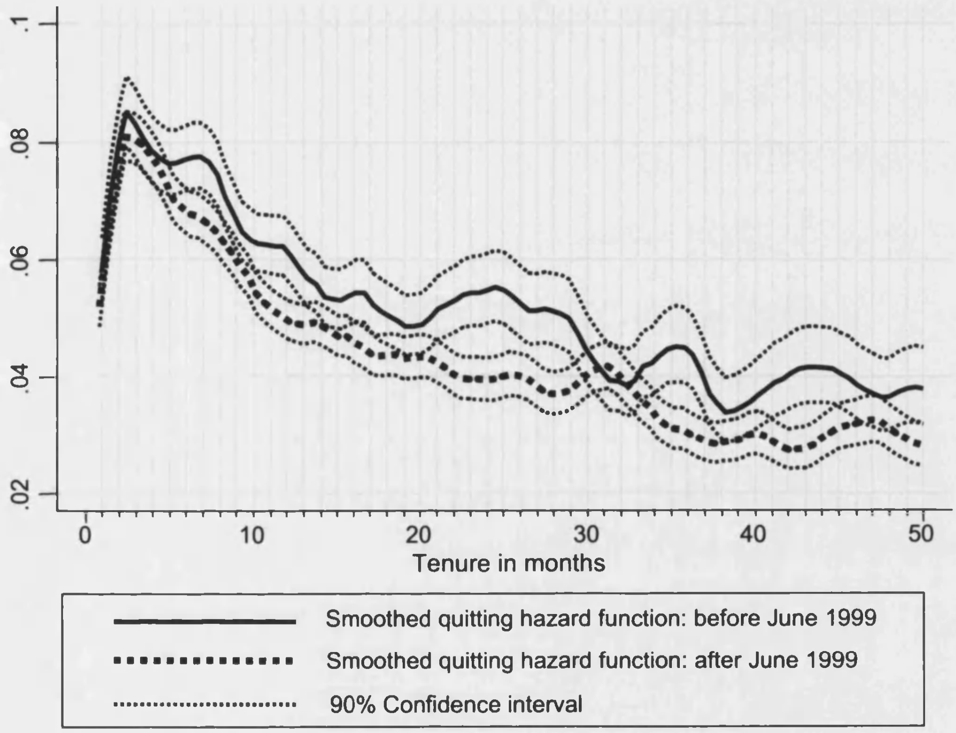
$$f_t(s) = \int_{s(\tau(t))}^{+\infty} f_{t-1}(s_1).p(s \mid s_1)ds_1$$

The firing hazard at tenure $t$ is then:

$$h(t)=\frac{\int_{-\infty}^{s(\tau(t))} f_t(s)ds}{\int_{-\infty}^{+\infty} f_t(s)ds}$$

APPENDIX 2: robustness checks

Figure 12: Kaplan-Meier estimates of the quit hazard before and after the reform




Notes: same as Figure 8, except that here the failure is quit.

# CHAPTER 2

# THE DETERMINANTS OF THE SEPARATION HAZARD IN A MODEL WITH LEARNING AND TIME-VARYING MATCH QUALITY

IOANA MARINESCU

ABSTRACT. People and organizations enter relationships, learn about them, adapt to them, and sometimes decide to leave them. This paper analyzes the impact of uncertainty, random shocks, time-varying separation costs and discount rates on the decisions of agents to separate from ongoing relationships. It also examines how such parameters affect the separation hazard, thus paving the ground for empirical analysis. If relationship quality is constant over time, the hazard converges to 0 as relationship length increases, whereas it converges to a positive value if relationship quality is subject to random shocks. In all models examined, higher separation costs and higher discount rates lower the separation threshold, i.e. they make agents more willing to continue with lower quality relationships; this in turn decreases the hazard of separation. A change in uncertainty or in the prevalence of random shocks has a negligible impact on the separation threshold, but a big impact on the separation hazard; moreover, the impact on the separation hazard depends on whether relationship quality changes over time. Overall, the model is very general and can allow us to understand and statistically analyze employment relationships, marriages, firm-supplier relationships, etc.

## 1. INTRODUCTION

When people enter a relationship, be it professional or personal, they usually do not know with certainty how good this relationship is for them. Moreover, a relationship that is good today may become undesirable tomorrow. Given this uncertainty, how do people and organizations decide whether to continue or separate from a relationship? In non-experimental empirical settings, we never observe the full information available to agents, but we are typically able to observe their separation decisions and determine how long the relationship was at the time of separation. These observations allow us to empirically estimate the separation hazard, i.e. the probability that a relationship is terminated given that it has survived so far. This paper develops a model which can ultimately allow the researcher to infer the tenor of the hidden information agents base their separation decision on. The model assumes that agents have an unbiased prior belief about the distribution of match qualities among potential partners, i.e. they know how likely they are to find a given match quality when sampling from the population of potential partners. Then, agents observe signals of relationship quality over time, and decide whether to separate or not based on their updated belief about quality, the costs of separation, and their discount factor. Three versions of the model are considered. In all versions, signals are normally distributed conditional on true match quality. In the first version, match quality only takes two values and is constant over time. In the second version, match quality is normally distributed and constant over time. Lastly, in the third and most general version, match quality is normally distributed and evolves over time according to an AR(1) process. I analyze the effect of belief-shaping parameters, the cost of separation and the discount factor on the threshold for separation (i.e. the match quality such that the agent is indifferent between continuing and separating) and on the hazard of separation. In all three models, separation costs and a lower discount factor (or higher discount rate) decrease the separation threshold, and thus the separation hazard in all cases: this is intuitive since both parameters diminish the returns of separating in order to look for a better option. The effects of parameters entering the belief are more complex. One striking feature of the results involving a change in the parameters governing belief formation is that, as long as the prior expected value of a relationship does not change, other parameters entering the belief formation have a very limited impact on the threshold of separation. The parameters entering the belief are however very important in shaping the separation hazard, and the way they affect the hazard depends on whether actual relationship quality is allowed to vary over time; the details of these effects are however too complex to discuss in this introduction.

The contribution of this paper to the literature is three-fold. First, while some general results about the separation hazard have been derived by Jovanovic(1979) in a model where match quality was assumed to be normal and constant over time, this paper directly computes the quantitative effect of various parameters on the hazard of separation. Second, we introduce time-varying separation costs and analyze the impact of such variation on the separation hazard. Third, we generalize the Jovanovic(1979) model by allowing match quality to evolve over time according to an AR(1) process. This extension to time-varying match quality is highly relevant empirically. Indeed, it seems unrealistic to assume that relationship quality cannot change, and estimated job separation hazards (Farber(1994),Marinescu(2006a)) or divorce hazards (Weiss-Willis(1997), Svarer(2004)) are inconsistent with a constant match quality since they do not decline to 0 with relationship length.

This paper is organized as follows. In section 2, I describe a very general model of the optimal separation decision and analyze the impact of parameters on the separation threshold and hazard in this context. However, it is not possible to derive quantitative estimates of the impact of parameters in the general case, and I therefore move on to more specific assumptions about the distribution of match quality in sections 3 to 5. Specifically, in section 3, I compute the impact of parameters assuming that match quality can only take two values and does not vary over time. In section 4, I assume that match quality is normally distributed and cannot change over time, while section 5 relaxes this latter assumption by allowing match quality to follow and AR(1) process. Section 6 discusses the limitations and implications of the results. Finally, section 7 concludes.

## 2. The optimal separation decision as a Partially Observable Markov Decision Process (POMDP)

The goal of this paper is to model the decision of an agent[1] to continue or separate from a relationship. The relationship links the agent with a partner. It is assumed that the agent entering a new relationship does not know the exact value of such a relationship. The quality of the relationship, or match quality, is what makes the relationship valuable to the agent, the benefits (monetary and others) of the relationship to the agent. The agent holds a prior belief about the distribution of quality in the population of partners that it encounters. Then, at each period, the agent observes a signal of the quality of the relationship. Based on these signals, the agent updates its belief using Bayes' rule, and decides whether to continue with the current relationship, or end it and start a new one[2]. If the agent decides to end the relationship, it has to pay a cost $f(k)$ which is a function of the length of the relationship $k$.

Formally, such a decision process is well modeled in the framework of Partially Observable Markov Decision Processes (see Hauskrecht(2002) for a full description of such models). This model allows solving for the optimal policy of the agent. The model is fully specified by states, actions, observation and transition functions, reward function, discount and planning horizon.

### 2.1. Definitions.

2.1.1. *States, actions.* The *state of the world* is defined by a vector of two variables: the length of the current relationship $k$, and the quality of the agent-partner match, $q$. The length of the relationship is perfectly observed

[1]The agent may be a person or an organization. I will use the pronoun 'it' to refer to the agent in this abstract general context.
[2]This means that there is no explicit account for search in this model, and agents cannot choose to stay unmatched. For a discussion of the implications of such assumptions, see the discussion in section 6.

by the agent. Moreover, to be realistic and simplify calculations, we assume that the length of the relationship is limited to some length $k_{max}$. As to match quality, it is assumed to take a finite number of values[3] . All these hypotheses imply together that the state space is finite.

Using information coming from previous experience or some other source, the agent forms an idea of how likely it is that, when forming a relationship, that relationship will be turn out to be of a given quality. The agent thus has a prior belief about match quality, which is defined by a prior distribution of qualities $P(q_0)$. I will denote by $\bar{q}$ the expected value of this prior distribution. In sections 3 to 5, I will examine specific hypotheses about the quality state space and the prior distribution over that space. I will denote by $q_k$ the value of match quality at length $k$, thus allowing it to be time-varying.

At every time step, the agent has two possible actions $a$. it can continue the current relationship ($a = C$) or separate from the current partner and begin a new relationship with another partner ($a = S$).

2.1.2. *Observation and transition functions.* The observation function gives, for each action and actual match quality, the probability of observing a given signal $z$, i.e. $P(z|a, q)$. Note that the observation function is assumed to be independent of the length of the relationship. I will denote by $z_k$ the value of the observation at length $k$. In this paper, the observation function will always be a Gaussian:

$$P(z|C, q) = P(z|S, q) = N(\bar{z}, \sigma_{obs}) \tag{2.1}$$

where $\bar{z}$ is the expected value of the observation. Note that $\bar{z}$ has to depend on match quality, otherwise it would not be an informative observation. In

[3]In some of the models used in this paper, match quality is assumed to take an infinite number of values. However, when computing the solution to the agent's problem, we will discretize that infinite space, and so the number of match qualities used in computations will remain finite.

what follows, I will explore several specifications of the relationship between $\bar{z}$ and actual match quality.

The transition function attributes a probability to each new possible state as a function of the current state and the agent's action. Before I define this function, a few remarks are in order about the notation. There is a perfect correspondence between the length of the relationship and the action taken, and so, to simplify notation, I will dispense with the specification of the action when the latter is evident given the length of the relationship. More precisely, a relationship length of 1 indicates a separation during the previous time period, and any $k > 1$ indicates a decision to continue the relationship at the previous period. Indeed, at the beginning of a given relationship, the length of the relationship gets reset to 0[4]. The decision about a new relationship (length 0) is by definition "continue", because in this case beginning a relationship and continuing it are one and the same thing. Thus, the length $x$ of the relationship evolves deterministically, so that if $x = k$ then at the next time step:

$$x = \begin{cases} k + 1 \text{ if } a = C \\ 0 \text{ if } a = S \end{cases}$$

We are now ready to specify the match quality transition function, i.e. the probability of a given match quality next period given current match quality. This entails specifying the initial distribution of match qualities, which in this case is the prior distribution $P(q_0)$, and then how match quality evolves over time starting from there. The evolution of match

[4]In the rest of the paper, I will often use length and period interchangeably, because as long as there is no separation, they are, up to a constant, the same.

quality is thus governed by the following equations:

$$P(q_1|q_k) = P(q_0) \tag{2.2}$$

$$P(q_{k+1}|q_k) = N(\Psi(q_k), \sigma_q) \tag{2.3}$$

where $P(q_0)$ is the prior distribution of match qualities. Note that I restrict $P(q_{k+1}|q_k)$ to normal distributions. $\Psi(q_k)$ denotes a deterministic function of $q_k$. This allows the modeling of deterministic drifts: relationships can indeed get better or worse over time. If $\Psi(q_k) = q_k$ and $\sigma_q = 0$, then match quality is constant over time. The transition function is such that the state of the world at time $t+1$ only depends on the action of the agent and the state of the world at time $t$, and not on the whole history of actions and states, i.e. the state process is Markovian.

2.1.3. *Belief and belief transition function.* A belief state is a distribution of probability over the states of the world. While the length of the relationship is known with certainty, the belief about match quality needs to be specified as a probability distribution. Given the prior distribution of match qualities, the transition function and the observation function, it is possible to us Bayes' rule and compute the belief as a function of the history of observations from length 0 to the current length $k$ of the relationship. Let $\hat{q}_k$ be the expected value of the belief distribution at length $k$. The belief distribution can be calculated recursively:

$$P(q_k|z_{1:k}) = \begin{cases} P(z_1|q_1)\sum_{q_0} P(q_1|q_0)P(q_0) & \text{if } k = 1 \\ \\ \frac{P(z_k|q_k)P(q_k|z_{1:k-1})}{P(z_k|z_{1:k-1})} & \text{if } k > 1 \end{cases} \tag{2.4}$$

$$\hat{q}_k = E(q_k|z_{1:k}) \tag{2.5}$$

where $P(q_0)$ is the prior distribution of match quality. $z_{1:k}$ is a shorthand for the set of observations from length 1 to length $k$, i.e. $\{z_1, ..., z_k\}$. $P(z_k|q_k)$

is the observation function. The belief transition function[5], $P(q_k|z_{1:k-1})$, attributes a probability to each possible belief as a function of the previous belief and the agent's action. It is defined as:

$$P(q_1|z_{1:k-1}) = P(q_1) = \sum_{q_0} P(q_1|q_0)P(q_0) \tag{2.6}$$

$$P(q_k|z_{1:k-1}) = \sum_{q_{k-1}} P(q_k|q_{k-1})P(q_{k-1}|z_{1:k-1}) \tag{2.7}$$

$P(q_{k-1}|z_{1:k-1})$ is the previous period belief distribution. Equation 2.6 corresponds to the case where the agent separates, and equation 2.7 to the case where the agent continues with the current relationship. Note that if the agent separates from the current relationship, the next belief state does not depend on the current belief state, but on the prior belief.

As match qualities are not directly observed, the agent has to rely on beliefs to make its decisions. As we have just seen, beliefs depend on the prior distribution, the set of observations, the observation function and the transition function. Beliefs thus fully summarize what the agent knows about the system. Therefore, it is convenient to express the agent's decision problem in the belief space. It is important to note here that the decision process expressed in terms of beliefs (rather than actual states of the world) is Markovian. That is, it can be shown that in a POMDP, the belief about the state of the world at time $t+1$ only depends on the action of the agent and the belief of the agent about the state of the world at time $t$, and not on the whole history of actions and beliefs (see for example Cassandra 1998).

2.1.4. *Reward, discount, horizon.* The reward function $R$ associates a reward to each possible combination of belief and action continue (C) or

[5]From now on, whenever I refer to "the transition function", I will mean the belief transition function, not the state transition function.

separate (S):

$$R_C(P(q_k|z_{1:k})) = \hat{q}_k \tag{2.8}$$

$$R_S(P(q_k|z_{1:k})) = \bar{q} - f(k) \tag{2.9}$$

where $f(k)$ is a separation cost which depends on the length of the relationship $k$, and it is assumed that The agent discounts the future at rate $\delta \in [0, 1]$.

The reward function can be derived from two possible hypotheses about the observability of the per period benefit of the relationship to the agent. Either the benefit is not directly observed but is known to be equal to the relationship quality, and to be realized after the observation: in this case, the benefit is trivially equal to the agent's belief. Or the benefit is equal to the observation next period: in this case, if we define the observation to have the same expected value as the actual quality, then $\hat{q}_k$ is indeed the agent's best estimate of the expected value of the observation, and hence the reward, at the next period.

The separation cost $f(k)$ covers the direct cost of ending the current relationship, such as a firing cost in the case of the employment relationship. It also covers the costs of beginning a new relationship, such as hiring costs. If the two partners have diverging interests over separation, i.e. if for example it is harder for the worker to find a new job than for the firm to find a new worker, then the model is not complete because it does not explicitly account for both partners' rewards. However, if these diverging interests are known ex ante and do not depend on match quality, then the party that is relatively more advantaged by the separation can agree ex ante to make a fixed payment to the other party. This case is covered by the model since the cost $f(k)$ can also include any such payments.

The definition of the reward function is compatible with a Nash bargaining solution where the two partners split the surplus, so that, while the relationship continues, each partner gets a fixed share. Suppose that $\alpha$ is the share received by the modeled agent. The reward of the agent would then be $\alpha\hat{q}_k$ if continuing and $\alpha\bar{q}$ if separating; but this change is not substantial since it simply amounts to rescaling the distribution of match quality.

The planning horizon of the agent is assumed to be infinite. This means that the agent is infinitely lived; or alternatively, the agent's retirement from the relationships market is at some date so far away in the future that given the discount factor, it does not play any role in the agent's current decisions[6].

2.1.5. *Value function.* We now need to define what it means for the agent to follow an optimal strategy. To do so, I will first define the notion of a strategy or policy, and the value function for a policy. In the context of this model, at each time step, the agent has the choice between two actions: continue or separate. The agent chooses one of these actions depending on its current belief and the length of the relationship $k$. Define a policy $\pi$, which gives for each belief and relationship length the action to be taken. Define the $Q$ function $Q^\pi(P(q_k|z_{1:k}), a)$ as the expected return of taking action $a$ today and then following the policy $\pi$ in the future. The value function $V^\pi(P(q_k|z_1 : z_k))$ gives the current and future rewards of the agent as a function of current belief, assuming that the agent follows policy $\pi$ from now on. The optimal policy maximizes $V^\pi(P(q_k|z_1 : z_k))$, and gives rise to the optimal value function $V^*(P(q_k|z_1 : z_k))$. The optimal action value

[6]The model is thus not quite adequate for explaining the behavior of "old" agents. That is typically not a problem if the agent considered is an organization, but may be relevant if the agent is a person.

function $Q^*$ is defined as a function of the optimal value $V^*(P(q_k|z_1:z_k))$:

$$Q^*(P(q_k|z_{1:k}),C) = \hat{q}_k + \delta \sum_{q_{k+1}} P(q_{k+1}|z_{1:k})V^*(P(q_{k+1}|z_{1:k+1})) \tag{2.10}$$

$$Q^*(P(q_k|z_{1:k}),S) = \bar{q} - f(k) + \delta \sum_{q_1} P(q_1)V^*(P(q_1|z_1)) \tag{2.11}$$

The optimal value is given by the Bellman equation:

$$V^*(P(q_k|z_{1:k})) = \max_{a\in[C,S]} Q^*(P(q_k|z_{1:k}),a) \tag{2.12}$$

Given the assumptions that I will be using about the distribution of match qualities and observations, the belief distribution at length $k$, $P(q_k|z_{1:k})$, can be summarized by $(\hat{q}_k, k)$, i.e. the expected value of the belief distribution and the length of the relationship. To represent all possible values of the belief distribution, we can thus use a $n_q \times k_{max}$ matrix, where $n_q$ is the number of possible values for the expected value of match quality and $k_max$ is the maximum possible length for a relationship. To simplify notation in what follows, I will use $\hat{q}_k$ to summarize the belief distribution, since this expression already contains $k$.

Thus, the $Q$ functions for the actions "continue" and "separate" can be rewritten as:

$$Q^*(\hat{q}_k,C) = \hat{q}_k + \delta \sum_{\hat{q}_{k+1}} P(\hat{q}_{k+1}|\hat{q}_k)V^*(\hat{q}_{k+1}) \tag{2.13}$$

$$\begin{aligned} Q^*(\hat{q}_k,S) &= \bar{q} - f(k) + \delta \sum_{\hat{q}_1} P(\hat{q}_1|\bar{q})V^*(\hat{q}_1) \\ &= V_{new} - f(k), \text{ where } V_{new} = \bar{q} + \delta \sum_{\hat{q}_1} P(\hat{q}_1|\bar{q})V^*(\hat{q}_1) \end{aligned} \tag{2.14}$$

In this framework, the optimal policy followed by the agent is uniquely defined by $\tau(k)$, the belief such that the agent is indifferent between continuing or separating from its partner at relationship length $k$. In other terms, the threshold for separation $\tau(k)$ is defined by the equalization[7] of $Q$ functions for the actions "continue" (equation (2.13)) and "separate" (equation (2.14)), i.e.:

(2.15)
$$\tau(k) + \delta \sum_{\hat{q}_{k+1}} P(\hat{q}_{k+1}|\tau(k))V^*(\hat{q}_{k+1}) = \bar{q} - f(k) + \delta \sum_{\hat{q}_1} P(\hat{q}_1|\bar{q})V^*(\hat{q}_1)$$

(2.16)
$$\Leftrightarrow \tau(k) - \bar{q} + f(k) + \delta[\sum_{\hat{q}_{k+1}} P(\hat{q}_{k+1}|\tau(k))V^*(\hat{q}_{k+1}) - \sum_{\hat{q}_1} P(\hat{q}_1|\bar{q})V^*(\hat{q}_1)] = 0$$

2.2. **Computing the value function and the optimal policy.** To compute the optimal policy, one starts at the highest possible relationship length, i.e. $k_{max}$. At that point, because relationships come to a final ending, the value of a relationship is exactly equal to the value of a new relationship, minus final separation costs, i.e. $V_{new} - f(k_{max})$.

The algorithm starts with giving $V_{new}$ some arbitrary value. Then, at length $k_{max} - 1$, $Q(\hat{q}_k, S)$ and $Q(\hat{q}_k, C)$ are computed using equations 2.14 and 2.13. The optimal policy at $k_{max} - 1$ is then given by equation 2.12. These calculations are repeated for $k_{max} - 2, k_{max} - 3, ....$

It is thus possible to recursively compute the value up to length 0. $V_{new}$ is then defined as the value of a relationship with length 0 and quality $\bar{q}$ (the expected value of the prior distribution). We start the loop over again until $V_{new}$ is numerically identical to its value in the previous iteration[8]. One

[7]When performing computations, we only consider a finite number of match qualities. Therefore there will typically be no belief that makes the agent *indifferent* between continuing and separating: rather, the optimal action will be "separate" for some belief and "continue" for the next higher belief. In practice, I defined as the threshold the minimum expected belief such that it is optimal for the agent to continue the relationship.

[8]This is a special case of the "value iteration" algorithm, which has been shown to converge to the solution of the Partially Observed Markov Decision Problem (see Hauskrecht(2002)). Note however that this algorithm is not the fastest possible to establish the optimal policy, because we compute the values for all possible beliefs, whereas it is clear that if for some belief it is optimal to separate, then for all beliefs with lower expected value, it is also optimal to separate. If computation time were a concern, one could therefore use a faster algorithm.

can thus determine the value function and optimal actions for all beliefs and relationship lengths.

**2.3. The impact of parameters on the optimal policy.** In order to assess the impact of parameters on the optimal policy, one needs to show how the condition defining the threshold in equation (2.16) is affected when parameters change. Define $J$ as the left-hand side of equation (2.16), i.e.:

$$J = \tau(k) - \bar{q} + f(k) + \delta[\sum_{\hat{q}_{k+1}} P(\hat{q}_{k+1}|\tau(k))V^*(\hat{q}_{k+1}) - \sum_{\hat{q}_1} P(\hat{q}_1|\bar{q})V^*(\hat{q}_1)]$$

Consider some parameter $x$: we are interested in the sign of $\frac{\partial\tau(k)}{\partial x}$. Given that the threshold is defined by $J = 0$, we can use properties of implicit functions to determine the sign of $\frac{\partial\tau(k)}{\partial x}$. It is a known result that, if $J(\tau(k), x) = 0$, then $\frac{\partial\tau(k)}{\partial x} = -\frac{\partial J/\partial x}{\partial J/\partial\tau(k)}$. It is easy to show that $J$ increases with $\tau(k)$, which implies that $\partial J/\partial\tau(k) > 0$. Hence, we have:

$$sign(\partial\tau(k)/\partial x) = -sign(\partial J/\partial x) \tag{2.17}$$

In the general case, it is not possible to determine $sign(\partial J/\partial x)$. It becomes feasable, however, if one uses a few approximations[9]. First, I show that, if $k$ is small, then $\hat{q}_k = \hat{q}_{k+1}$ implies that $V^*(\hat{q}_k) \approx V^*(\hat{q}_k)$. This allows me to drop $k$ from $\hat{q}_k$. I then proceed to calculate the sign of $\partial J/\partial x$, which will involve some more approximations.

Early in the relationship, if $\hat{q}_k = \hat{q}_{k+1}$, there is only a negligible difference between $V^*(\hat{q}_k)$ and $V^*(\hat{q}_{k+1})$. This is because, given the existence of a discount factor, the maximum possible length $k_{max}$ is too far away in the future to influence the current value. Thus, at short relationship lengths, the value of a given belief does not change with relationship length $k$.

[9]For each of these approximations, I will explain why it may be correct. Moreover, all the approximations used are indeed good approximations in the models for which I explicitly compute the threshold in sections 3 to 5.

Using the approximation $V^*(\hat{q}_k) = V^*(\hat{q}_{k+1})$ for $\hat{q}_k = \hat{q}_{k+1}$ allows us to drop the subscript of $\hat{q}$, and the condition defining the threshold can then be rewritten as:

$$\tau(k) - \bar{q} + f(k) + \delta \sum_{\hat{q}} [P(\hat{q}|\tau(k)) - P(\hat{q}|\bar{q})] V^*(\hat{q}) = 0 \tag{2.18}$$

The function $J$ is consequently redefined as:

$$J = \tau(k) - \bar{q} + f(k) + \delta \sum_{\hat{q}} [P(\hat{q}|\tau(k)) - P(\hat{q}|\bar{q})] V^*(\hat{q}) \tag{2.19}$$

I now show that typically $\sum_{\hat{q}}[P(\hat{q}|\tau(k)) - P(\hat{q}|\bar{q})]V^*(\hat{q}) < 0$, which will be important in determining the sign of the derivative of $J$ with respect to some parameter $x$. Let $P_1(\hat{q})$ be a shorthand for $P(\hat{q}|\tau(k))$ and $P_2(\hat{q})$ a shorthand for $P(\hat{q}|\bar{q})$. The reasoning will rely on the fact that $P_1$ and $P_2$ are normal, $E(P_1) < E(P_2)$ and $V^*(\hat{q})$ increases in $\hat{q}$. The latter fact is trivial. Now I show that $E(P_1) < E(P_2)$. If $\tau(k) = \bar{q}$, then equation (2.18) implies that $f(k) = 0$. If $f(k) > 0$, then $\tau(k) < \bar{q}$, i.e. in the presence of positive separation costs, the threshold for separation is lower than the expected quality of a new match. If $\Psi(q_k) = q_k$ in equation (2.3), $E(P_1) = \tau(k)$ and $E(P_2) = \bar{q}$, and therefore, $E(P_1) < E(P_2)$. If $P_1$ and $P_2$ have the same variance, then, as $E(P_1) < E(P_2)$ and $V^*(\hat{q})$ increases in $\hat{q}$, we have that indeed $\sum_{\hat{q}}[P(\hat{q}|\tau(k)) - P(\hat{q}|\bar{q})]V^*(\hat{q}) < 0$. Indeed, the expression $\sum_{\hat{q}}[P(\hat{q}|\tau(k)) - P(\hat{q}|\bar{q})]V^*(\hat{q})$ is the difference of two weighted means of the increasing series $V(\hat{q})$: in this case, the mean around $\tau(k)$ is smaller than the mean around $\hat{q}$.

More generally, the variances of $P_1$ and $P_2$ may differ and so the means $\sum_{\hat{q}} P(\hat{q}|\tau(k))V^*(\hat{q})$ and $\sum_{\hat{q}} P(\hat{q}|\bar{q})V^*(\hat{q})$ use different weights. Denote by $\sigma_1$ the standard deviation of $P_1(\hat{q})$ and $\sigma_2$ the standard deviation of $P_2(\hat{q})$. As long as $\sigma_1$ is not much greater than $\sigma_2$, we still have $\sum_{\hat{q}}[P(\hat{q}|\tau(k)) -$

$P(\hat{q}|\bar{q})]V^*(\hat{q}) < 0$. In almost all cases that we will examine, $\sigma_1 < \sigma_2$[10] , and so we typically have $\sum_{\hat{q}}[P(\hat{q}|\tau(k)) - P(\hat{q}|\bar{q})]V^*(\hat{q}) < 0$.

2.3.1. *Impact of the separation cost on the optimal policy.* Taking the derivative of $J$ with respect to $f(k)$, we get:

$$\frac{\partial J}{\partial f(k)} = 1 + \delta \sum_{\hat{q}} (P_1(\hat{q}) - P_2(\hat{q})) \frac{\partial V^*(\hat{q})}{\partial f(k)} \tag{2.20}$$

$\partial V^*(\hat{q})/\partial f(k)$ decreases with $\hat{q}$ because higher quality matches have a lower probability of being eventually dissolved and thus the agent is less likely to bear the firing cost for higher quality matches. Since $\partial V^*(\hat{q})/\partial f(k)$ decreases with $\hat{q}$, the second term of equation (2.20) is positive[11]. Thus, $\partial J/\partial f(k) > 0$, which implies that $\partial \tau(k)/\partial f(k) < 0$. As is intuitive, this means that higher separation costs make the agent more willing to pursue relationships of lower value. Hence, we have:

**Proposition 1.** *Higher separation costs $f(k)$ lower the threshold for separation $\tau(k)$*[12]

This implies in particular that if the separation cost increases over time, then the threshold decreases over time. For example, if there is some sort of probation period, with a low constant separation cost followed by a higher constant separation cost, then at the period when the separation cost increases, the threshold will decrease. Moreover, in such a case, the threshold will slightly increase at the end of the probation period (i.e. the initial period with low separation cost). This is explained by the following

[10]This is because it is typically the case that the agent's belief gets more precise over time, and so since $P(\hat{q}|\bar{q})$ is taken at length 0, its variance is greater than the variance of $P(\hat{q}|\tau(k))$, which is taken at some length at least equal to 1.

[11]This is for the same reasons why the fact that $V^*(\hat{q})$ increases in $\hat{q}$ implies that $\sum_{\hat{q}}[P(\hat{q}|\tau(k)) - P(\hat{q}|\bar{q})]V^*(\hat{q}) < 0$.

[12]Most propositions in this paper (including this one) are dependent on the approximations used. However, they are still useful to understand the logic of the model, and they relate to each other in such a way that it is useful to number them. When a proposition depends on approximations, I will signal it in a footnote.

consideration. As the end of the probation period approaches, the value of separating from the relationship stays the same but the value of continuing the relationship decreases because of the increased probability that a bad quality relationship will have to be terminated under the higher post-probation separation cost. Because the value of continuation decreases relative to the value of separation as the agent approaches the end of the probation period, the threshold increases.

2.3.2. *The impact of the discount factor on the threshold of separation.* Taking the derivative of $J$ with respect to $\delta$, we get:

$$\frac{\partial J}{\partial \delta} = \sum_{\hat{q}} [P_1(\hat{q}) - P_2(\hat{q})][\delta \frac{\partial V^*(\hat{q})}{\partial \delta} + V^*(\hat{q})] \tag{2.21}$$

It seems plausible that $\partial V^*(\hat{q})/\partial \delta$ is roughly constant over $\hat{q}$, as a lower discount factor roughly proportionally reduces the value of all levels of match quality[13]. If this assumption is valid, then $\sum_{\hat{q}}[P_1(\hat{q}) - P_2(\hat{q})]\delta \frac{\partial V^*(\hat{q})}{\partial \delta} = 0$: this is because both $P_1(\hat{q})$ and $P_2(\hat{q})$ add up to 1 over $\hat{q}$. Then, equation (2.21) simplifies to:

$$\frac{\partial J}{\partial \delta} = \sum_{\hat{q}} [P_1(\hat{q}) - P_2(\hat{q})]V^*(\hat{q}) \tag{2.22}$$

We have already established that the above expression is negative. Thus, from equation (2.17), we infer that:

**Proposition 2.** *The threshold of separation $\tau(k)$ increases with a higher discount rate $\delta$*[14].

2.3.3. *The impact a change in the transition function on the threshold of separation.* We are now interested in the effect of some parameter $x$ that enters both $P_1$ and $P_2$. The derivative of $J$ with respect to such a parameter

[13]This is verified in all specific cases I will analyze in sections 3 to 5.

[14]This proposition depends on the approximations used.

is:

$$(2.23) \qquad \frac{\partial J}{\partial x} = \sum_{\hat{q}} (\frac{\partial P_1(\hat{q})}{\partial x} - \frac{\partial P_2(\hat{q})}{\partial x}) V^*(\hat{q}) + (P_1(\hat{q}) - P_2(\hat{q})) \frac{\partial V^*(\hat{q})}{\partial x}$$

As previously, if $\frac{\partial V^*(\hat{q})}{\partial x}$ is constant, then the second term drops out. Assume that indeed the second term drops out.

If separation costs are very small, then $\tau(k) \approx \bar{q}$, i.e. $P_1(\hat{q}) \approx P_2(\hat{q})$; thus, $\frac{\partial P_1(\hat{q})}{\partial x} - \frac{\partial P_2(\hat{q})}{\partial x}$ is very close to 0, implying that $\frac{\partial J}{\partial x} \approx 0$, i.e. a change in the transition function has no effect on the threshold.

If separation costs are bigger, then one can rely on other approximations. $P_1$ and $P_2$ are probability distributions, so we must always have $\sum_{\hat{q}} P_i(\hat{q}) = 1, i = 1, 2$, therefore, $\sum_{\hat{q}} \frac{\partial P_i(\hat{q})}{\partial x} = 0, i = 1, 2$, and so $\sum_{\hat{q}} (\frac{\partial P_1(\hat{q})}{\partial x} - \frac{\partial P_2(\hat{q})}{\partial x}) = 0$. Moreover, $P_1$ and $P_2$ are normal, so they are close to 0 for values of $\hat{q}$ far away from their respective means. Any change in the variance of these distributions will mainly have an impact in the neighborhood of their means. Therefore, $\frac{\partial P_1(\hat{q})}{\partial x} - \frac{\partial P_2(\hat{q})}{\partial x}$ is only significantly different from 0 in the neighborhood of $\tau(k)$ and $\bar{q}$, i.e. in the neighborhood of the separation threshold. But in the neighborhood of the separation threshold, $V^*(\hat{q})$ is almost constant since it is equal to the constant $Q^*(\hat{q}, S)$ below the threshold, and increases slowly above the threshold. If $V^*(\hat{q})$ can be assumed to be constant in the neighborhood of $\tau(k)$ and $\bar{q}$, then the first term of equation (2.23) is 0. In that case, we again obtain that $\frac{\partial J}{\partial x} \approx 0$, so that a change in the variance of the transition function does not affect the threshold for separation. While the conclusion above relies on quite a few approximations, it does indicate that, in most cases, a change in the variance of the transition function will have a very limited impact on the threshold of separation.

2.4. **Hazard of separation.** Deriving the impact of parameters on the hazard of separation is an important task because the separation hazard can be computed from empirical data, while the threshold for separation is typically not observed. The theoretical hazard of separation is the result of infinitely many agents confronted with the same separation problem; it summarizes the average separation behavior of agents over relationship lengths.

One can compute the theoretical separation hazard once the threshold for separation is known. Note that at length 0, when no observation has been made yet, $\hat{q}_0 = \bar{q}$ for all matches, i.e. for all agents the belief is the same as the prior. Let $p_k(\hat{q}_k)$ be the density of agents who hold a belief with mean $\hat{q}_k$ at length $k$, given that they follow the optimal policy embodied in $\tau(k)$. The initial values for the distribution of agents' expected beliefs about match quality are:

$$p_0(\hat{q}_0) = \begin{cases} 1 \text{ if } \hat{q}_0 = \bar{q} \\ 0 \text{ otherwise} \end{cases} \tag{2.24}$$

$$p_1(\hat{q}_1) = \sum_{\hat{q}_0} p_0(\hat{q}_0) P(\hat{q}_1|\hat{q}_0) = P(\hat{q}_1|\bar{q}) \tag{2.25}$$

The hazard of separation at length k, $h_k$, can then be computed recursively, starting at $k = 1$:

$$h_k = \sum_{\hat{q}_k=\hat{q}_{min}}^{\hat{q}_k=\tau(k)} p_k(\hat{q}_k) \tag{2.26}$$

$$p_k(\hat{q}_k) = 0 \text{ if } \hat{q}_k \leq \tau(k) \tag{2.27}$$

$$p_k(\hat{q}_k) = \frac{p_k(\hat{q}_k)}{\sum p_k(\hat{q}_k)} \tag{2.28}$$

$$p_{k+1}(\hat{q}_{k+1}) = \sum_{\hat{q}_k} p_k(\hat{q}_k) P(\hat{q}_{k+1}|\hat{q}_k) \tag{2.29}$$

Equation (2.28) insures that the mass of agents is always normalized to 1. $P(\hat{q}_{k+1}|\hat{q}_k)$ can be computed using the belief transition function defined in equation (2.7). Note that if match quality does not change over time, then the hazard declines to 0 as relationship length $k$ increases: indeed, over time, the knowledge of the agent becomes more and more precise, and eventually there are no more matches to terminate as only those that are above the threshold with certainty remain.

2.5. **The impact of parameters on the hazard of separation.** Parameters affect the hazard of separation through their effect on the threshold and the transition function. Therefore, we will first discuss the impact of changes in the threshold and the transition function on the hazard, and then proceed to the full analysis of the impact of parameters.

2.5.1. *Impact of the threshold and the transition function on the hazard of separation.* For a given distribution $p_k(\hat{q}_k)$, the effects of a change in the threshold of separation $\tau(k)$ or the transition function $P(\hat{q}_k|\hat{q}_{k-1})$ on the hazard of separation $h_k$ are straightforwardly described. The impact of the transition function on the hazard of separation for a given threshold is defined by equations (2.25) and (2.29).

**Proposition 3.** *For a given threshold $\tau(k)$ and a given distribution of continuing relationships $p_k(\hat{q}_k)$, a higher variance for the transition probability $P(\hat{q}_k|\hat{q}_{k-1})$ implies a higher separation hazard $h_k$.*

Indeed, a higher variance for the transition probability $P(\hat{q}_k|\hat{q}_{k-1})$ implies that more matches will cross the threshold from one period $k$ to the next (see equation 2.29). Thus, any change in parameters that increases the standard error of the transition function $P(\hat{q}_k|\hat{q}_{k-1})$, will, for a given policy and distribution of continuing relationships, increase the hazard of separation at length $k$.

For a given transition function, the threshold for separation affects the hazard of separation as described by equation (2.26).

**Proposition 4.** *The higher the threshold $\tau(k)$, the higher the hazard of separation $h_k$.*

Indeed, a higher threshold implies that a higher expected value of the belief distribution is needed for the agent to continue the relationship. Describing the effects of a change in the threshold of separation or the transition function on the hazard of separation all other things equal is a useful first step.

However, if one wants to assess the impact of a change in parameters on the hazard of separation derived from an optimal policy, it is necessary to examine the effect of each parameter on both the the transition function and the threshold.

2.5.2. *Impact of parameters entering the transition function on the separation hazard.* If a parameter affects the variance of the transition function but has little impact on the threshold, then proposition 3 determines the impact of such a parameter. If, however, the parameter has a significant impact on both the threshold and the transition function, the effect on the hazard function cannot be predicted but has to be calculated numerically. Let us now examine the impact of the first type of parameters: can we say anything about it since proposition 3 is conditional on the distribution of continuing relationships $p_k(\hat{q}_k)$?

First, note that no change in the variance of the transition function can affect the starting point for the calculation hazard: indeed, from equation (2.24), $p_0(\hat{q}_0)$ only depends on $\bar{q}$. Second, one should realize that the hazard of separation at length $k$ is determined by the successive application of the transition function to the initial distribution $p_0(\hat{q}_0)$, with a truncation of the distribution below the threshold at each step. Thus, if a parameter increases the variance of the transition function at all lengths and does not affect the threshold, then, from Proposition 3, it increases the hazard of separation at length 1 for sure. However, this effect may be reversed with increasing length, depending on exactly how the variance of the transition function evolves with length[15]. Thus, we have that:

**Proposition 5.** *A parameter that increases the variance of the transition function at short lengths, and does not affect the threshold, increases the hazard of separation at short lengths.*

2.5.3. *Impact of separation costs on the separation hazard.* Higher separation costs decrease the threshold and do not influence the transition function, therefore, from proposition 4:

**Proposition 6.** *Higher separation costs $f(k)$ decrease the separation hazard $h_k$ for all lengths $k$[16].*

If separation costs do not depend on $k$, then higher separation costs decrease the hazard of separation at all lengths. If there is a probation period, then the hazard of separation will be higher during the probation period relative to the post-probation period. Moreover, because the threshold increases when approaching the separation threshold, the hazard of separation also increases, creating a spike at the end of the probation period.

2.5.4. *Impact of the discount factor on the separation hazard.* A higher discount rate increases the separation threshold and does not affect the transition function, therefore, from proposition 4:

**Proposition 7.** *A higher discount factor $\delta$ increases the separation hazard $h_k$ for all lengths $k$[17].*

[15]This point will become clearer in section 5, where such a reversal is observed under some parameter values.

[16]This proposition depends on the approximations used in assessing the effect of separation costs on the threshold of separation.

[17]This proposition depends on the approximations used in assessing the effect of the discount rate on the threshold of separation.

We have now completed the exploration of what can be said about the effects of parameters on the threshold of separation and the separation hazard under the assumption that the transition function is normal. In the coming sections, I will compute and analyze these effects under more specific assumptions about the distribution of match quality and the observation function.

## 3. Two quality levels

**3.1. Model specification.** In this section, I assume that match quality can be either good or bad. More specifically a good match has a per-period value of 1 and a bad match a value of 0. I assume that a proportion g of the matches is good whereas a proportion 1-g is bad. Therefore, the expected value of the prior distribution is:

$$\bar{q} = 1g + 0(1-g) = g \tag{3.1}$$

At each period, the agent observes a normally distributed signal about the quality of the match. The signal for a good match is normally distributed with mean 1 and variance $\sigma^2_{obs}$, whereas for a bad match it is normally distributed with mean -1 and variance $\sigma^2_{obs}$. The belief of the agent that the match is good can be written $b(s_k)$ where $s_k$ is the sum of all signals observed up to length $k$[18]. Because there are only two values of match quality, the belief that the match is bad is $1 - b(s_k)$. The expected value of $s_k$ given the quality of the match is described by a normal distribution. Let $\varphi_g(s_k)$ be the probability of getting a sum $s_k$ of observations by length $k$ when the true match quality is good; $\varphi_b(s_k)$ denotes the same probability when the match is bad. $\varphi_g(s_k)$ is normally distributed with

[18]Given that match quality is fixed over time for a given match, and that the observation function does not depend on relationship length (no observation is more precise than another), the sum of observations is a sufficient statistic for the full history of observations.

mean $k$ and variance $k\sigma_{obs}^2$. Symmetrically, $\varphi_b(s_k)$ is normal with mean $-k$ and variance $k\sigma_{obs}^2$. Using Bayes' rule we can then compute all possible beliefs. We have:

$$b(s_k) = \frac{g\varphi_g(s_k)}{g\varphi_g(s_k) + (1-g)\varphi_b(s_k)} \tag{3.2}$$

$$= \frac{g\exp(\frac{s}{\sigma_{obs}^2})}{g\exp(\frac{s}{\sigma_{obs}^2}) + (1-g)\exp(\frac{-s}{\sigma_{obs}^2})} \tag{3.3}$$

Note that after simplifying the length $k$ drops out of the belief, so that the agent's belief only depends on the sum of observations. Finally, one needs to specify the belief state transition function.

$$P(b(s_k)|s_{k-1}), C) = b(s_k)N(s_k - s_{k-1}, 1, \sigma_{obs}) + (1-b(s_k))N(s_k - s_{k-1}, -1, \sigma_{obs}) \tag{3.4}$$

$$P(b(s_k)|s_{k-1}), S) = g \tag{3.5}$$

where $N(a, b, c)$ stands for the normal distribution with mean $b$ and standard deviation $c$, evaluated at $a$. Since $s_{k-1}$ is the sum of observations at length $k-1$, $s_k - s_{k-1}$ is the value of the observation made at length $k$. Equation (3.4) says that if the agent continues with the current relationship, then with a probability $b(s_k)$ the relationship is good, and so the mean of the observation at the next period is 1, and with a probability $1 - b(s_k)$ the match is bad and so the mean of the observation at the next period is -1.

3.2. **Belief space discretization.** As we have just seen, for the purpose of value calculation, belief is fully summarized by the value $b(s_k) \in [0, 1]$. For computational purposes, the interval $[0, 1]$ is divided in discrete steps. It would seem natural to divide it in some $n$ equal steps. However, such a division would lead to computational problems. Indeed, as the agent's knowledge of match quality gets more precise with more observations, beliefs will tend to get very close to 0 if the match is bad, and to 1 if the match is good. This means that, to adequately represent the evolution of beliefs over time, the disrectization needs to be much more precise, i.e. to have smaller steps, in the neighborhood of 0 and 1. To achieve this, I choose to divide the interval $[0, 1]$ in $n$ values, but with varying intervals in between values: specifically, the values of match qualities are defined by $normcdf(X, 0.5, 0.1)$, where $X$ is a vector of $n$ equally spaced values between 0 and 1, and $normcdf(X, 0.5, 0.1)$ is the cumulative distribution function, evaluated at point $X$, of a normal probability distribution with mean 0.5 and standard deviation 0.1. The mean of 0.5 is chosen so that the step size is symmetric over the $[0, 1]$ interval, and the standard error was chosen inductively to be very small (so that there are indeed many more values in the neighborhood of 0 and 1), but still provide a reasonable coverage for the middle of the $[0, 1]$ interval. Given this discretization issue, it is important remember that quantities computed are only approximate, and especially so at long relationship durations[19].

3.3. **Results.** The parameters given in Figure 1 were used for the reference case. In all cases, the parameters in the reference case were chosen such that the resulting hazard function is similar to the empirical hazard of a worker getting fired in the United Kingdom in the late 1990's[20]. The crucial point is that parameters were chosen de facto to be such that the hazard of separation first increases and then decreases with relationship

[19]It is possible to make the computation more precise in this case by choosing to describe the problem in the space of sum of observations $s_k$ rather than beliefs. This is what I have done in my paper Marinescu(2006a). In the context of this paper, however, to keep computations more comparable across different hypothesis, I decided to keep the computations in the belief space.
[20]This empirical hazard was estimated in Marinescu(2006).

length, a pattern that is typically found in studies of the firing hazard (Farber(1994),Marinescu(2006a)) or the divorce hazard (Weiss-Willis(1997), Svarer(2004)).

For each parameter, I choose a few values below and above the reference value, and I compute the variance of the transition function, the separation threshold, and the resulting hazard of separation.

[Figure 1 about here.]

Because the transition function, threshold for separation and separation hazard for the reference case are plotted in the figures that follow, I will not plot them again here.

3.3.1. *The impact of separation costs.* As already pointed out, separation costs have no impact on the transition function. Figure 2 illustrates how higher separation costs do indeed lower the separation threshold. Note, moreover, that the threshold is constant with length in the beginning of the relationship, and thereafter it increases very slightly as the relationship approaches the maximal length; this pattern is preserved for almost all parameter values, and I shall signal when this is not the case. Since a higher separation cost decreases the threshold, it lowers the separation hazard (Figure 3[21]). Note that, as predicted, since match quality does not change over time, the hazard declines to 0 as relationship length increases.

[Figure 2 about here.]

[Figure 3 about here.]

One can also examine the impact of a probation period, i.e. instead of having constant separation costs over the length of the relationship, separation costs are constant in the beginning of the relationship, and then they increase to a higher and constant level after some given length. In

[21]From the figure, the hazard seems to be higher at higher tenures for higher firing costs; hazards for different separation costs in fact converge as relationship length increases, and the observed effect is due to a discretization artifact.

this case, I use a separation cost of 1.5 in the beginning, and 2.3 after the end of the probation period. I also show results for two different lengths of the probation period, that is 12 and 24 periods. The separation thresholds are plotted in Figure 4. As predicted, one observes an increase in the threshold before the end of the probation period, and lower thresholds afterwards. The hazards are plotted in Figure 5. As a result of the variations in the thresholds, the hazards increase right before the end of the probation period, producing spikes in separations. The spike is higher with a shorter probation period because at lower length there are more relatively low quality matches that are close to the threshold and have not yet been terminated.

[Figure 4 about here.]

[Figure 5 about here.]

3.3.2. *The impact of the discount factor.* As predicted, the threshold for separation increases with a higher discount factor (Figure 6). This implies that the hazard of separation increases, as shown by Figure 7[22].

[Figure 6 about here.]

[Figure 7 about here.]

Thus, the discount factor has virtually the same effect on the threshold and the firing hazard as the separation cost: both a higher separation cost and a lower discount factor shift the threshold for separation downwards in a quasi parallel fashion, inducing a decrease in the firing hazard which is more pronounced for shorter relationships. This means that, in empirical applications, if neither the separation cost nor the discount factor are observed, one cannot distinguish the effect of those two parameters by observing either the separation threshold or separation hazard. Only the observation of the value function itself would allow to distinguish the

[22]As for firing costs, the fact that the hazard seems to be higher at higher tenures for lower discount factors is due to a discretization artifact.

two: indeed, the discount factor affects all match qualities equally while the separation cost mostly affects relationships that are close to the threshold. Since the value function will typically not be observed, in most applications one would have to have some other source of knowledge about the discount factor or the separation cost. For example, it should often be reasonable to assume that the discount factor is stable over time, while it may vary across agents.

3.3.3. *The impact of the observation variance.* The observation variance affecting both the transition function and the threshold, both of these effects need to be taken into account in order to understand the impact on the separation hazard. Figure 8 plots the transition function for different observation standard deviations, and for belief 0.5. The transition function has the same standard deviation for all beliefs; only the mean changes. We can see that a higher observation standard deviation translates into a lower standard deviation for the transition function, which can be derived from equation (3.4). Intuitively, this is because a higher standard deviation of the observation makes it optimal for the agent to rely relatively more on the prior belief and less on the observation. If, as predicted, the threshold of separation does not change much with the change in the transition function, then the hazard of separation at low lengths decreases with a higher observation variance.

[Figure 8 about here.]

The effect of a higher observation standard deviation on the threshold is depicted in Figure 9: the threshold increases with a higher standard deviation of the observation. This increase is limited, but has an effect on the hazard of separation which is opposite to the direct effect of the transition function, and thus the total effect is undetermined.

[Figure 9 about here.]

As shown in Figure 10, the effect on the transition function dominates in the sense that the hazard at short lengths is lower with a higher standard deviation of the observation.

[Figure 10 about here.]

We have thus confirmed our predictions about the impact of separation costs and the discount factor on the separation hazard, and shown how the observation variance affects the hazard. We know turn to a more general specification of the problem, assuming that match qualities, instead of only taking two values, are normally distributed.

## 4. Constant quality, normally distributed

**4.1. Model specification.** In this section, it is assumed that match quality is normally distributed. Let every relationship be characterized by a true quality q that is drawn from $N(\bar{q}, \sigma_q)$.

In order to compute the belief transition function, one needs to specify how observations arise and how beliefs evolve over time as a result.

Let each observation $z_k$ about a given relationship be a noisy observation of the relationship's true quality $q$ drawn from $N(q, \sigma_{obs})$.

$$z_k = q + \epsilon_k^z \tag{4.1}$$

As the likelihood function is a Gaussian and the prior is Gaussian as well, we obtain a Gaussian posterior function:

$$P(q|z_{1:k}) = N(\hat{q}_k, \sigma_k) \tag{4.2}$$

where $z_{1:k}$ stands for the set of observations from length 1 to length $k$. We observe that $\sigma_k$ does not depend on any of the observations but only depends on the number of observations that are available. The definition of the variance of the belief in this case is well known:

$$\sigma_k^2 = \frac{\sigma_{obs}^2 \sigma_q^2}{t\sigma_q^2 + \sigma_{obs}^2} \tag{4.3}$$

To calculate the expected value for the quality $q$ after an additional observation $z_{k+1}$, we use the fact that we can analytically treat the multiplication of two gaussians:

$$\hat{q}_{k+1} = \alpha_k \hat{q}_k + (1 - \alpha_k) z_{k+1} \tag{4.4}$$

where $\alpha_k = \sigma_{obs}^2 / (\sigma_{obs}^2 + \sigma_k^2)$. With this recursion rule, it is possible to calculate the belief after a number of observations. We can also express this as a function of the mean $\bar{q}$ of the prior distribution and the sum of observations up to time $t$:

$$\hat{q}_k = \frac{\sigma_{obs}^2 \bar{q} + \sigma_q^2 \sum_{i=1}^k z_i}{t\sigma_q^2 + \sigma_{obs}^2} \tag{4.5}$$

Given these assumptions, a belief is fully specified by $\hat{q}_k$ and $\sigma_k$. The belief transition function will thus specify the probability of transition from a belief defined by $\hat{q}_k$, $\sigma_k$ to a belief defined by $\hat{q}_{k+1}$ and $\sigma_{k+1}$[23].

We now proceed to calculate the belief transition function. We have already given above the equation for $\sigma_k$ (equation 4.3). The expected quality $\hat{q}_k$, depends on the observations made (equation 4.4). On top of these equations, it will also be necessary to calculate how likely any given observation will be at the next step $k+1$ given our current knowledge $\hat{q}_k$: in other terms, we need to specify the transition function. The variance of the estimate of the quality at length $k$ is $\sigma_k^2$. The variance of the observation is $\sigma_{obs}$. As these two effects are additive and we know that variances add

[23] I will omit conditioning on $\sigma_k$ to simplify the notation. Indeed, $\hat{q}_k$ already contains $k$ and if $k$ is known, then $\sigma_k$ is immediately given by equation (4.3).

linearly we obtain the following equation:

$$P(z_{k+1}|\hat{q}_k, \sigma_k) = N(z_{k+1}, \hat{q}_k, \sqrt{\sigma_k^2 + \sigma_{obs}^2}) \tag{4.6}$$

where $N(a, b, c)$ stands for the normal distribution with mean $b$ and standard deviation $c$, evaluated at $a$. This probability can be equivalently expressed in the following way:

$$\begin{aligned} P(z_{k+1}|\hat{q}_k, \sigma_k) &= \int_{-\infty}^{+\infty} P(z_{k+1}|q)P(q|z_{1:k})dq \\ &= \int_{-\infty}^{+\infty} N(z_{k+1}, q, \sigma_{obs})N(q, \hat{q}_k, \sigma_k)dq \end{aligned} \tag{4.7}$$

As I chose to perform computations in the belief space, I need to express everything in terms of the mean of belief $\hat{q}_k$, and to eliminate the observations $z_k$ from the equations. This can be achieved by expressing $z_{k+1}$ as a function of the other variables using (4.4).

$$z_{k+1} = \frac{\sigma_{obs}^2}{\sigma_k^2}(\hat{q}_{k+1} - \hat{q}_k) + \hat{q}_{k+1} \tag{4.8}$$

Thus, because computations are performed in the belief space and not the observation space, the expression $\sigma_k^2 + \sigma_{obs}^2$ is not exactly the variance of the transition function. The variance of the transition function at length $k$, $VT_k$, can instead be calculated using the definition of the variance for a probability distribution with a known mean:

$$VT_k = \int_{-\infty}^{+\infty} (\hat{q}_{k+1} - \hat{q}_k)^2 P[z_{k+1}(\hat{q}_{k+1})|\hat{q}_k]d\hat{q}_{k+1} \tag{4.9}$$

4.2. **Belief space discretization.** Note that the specific value of $\bar{q}$ is not substantial for the calculations since the normal distribution is symmetric and defined over $\mathbb{R}$. $\bar{q}$ only matters relative to the separation cost. The discretization uses equally spaced values between some minimal and some maximal value of $\hat{q}_k$. To make the interpretation of $\hat{q}_k$ more intuitive, I chose the minimal value of $\hat{q}_k$ to be 0 and its maximal value to be $2\bar{q}$. Thus a positive separation cost is commensurate with the per period value of the relationship.

4.3. **Results.** The parameters used as a reference for calculations can be found in Figure 1. Note that, contrary to the 2-qualities case, the transition function here depends on the length of the relationship. Namely, as the relationship gets longer, the variance of the transition function decreases, as implied by equations (4.3) and (4.6). Therefore, when looking at the effect of parameters that affect the transition function, we need to document how the variance of the transition function is affected. Another slight difference with the 2-qualities case is that the threshold for separation increases a bit more with length $k$, even though this increase remains very limited (see for example Figure 11).

4.3.1. *The impact of separation costs.* An increase in separation costs has qualitatively the same effect as in the 2-qualities case, both on the threshold for separation (Figure 11) and on the separation hazard (Figure 12). Note that the separation hazard at lengths greater than 15 would look the same in the 2-qualities case if it were not for the discretization artifact already mentioned. Again, since match quality does not change over time, hazards converge towards 0 at high lengths.

[Figure 11 about here.]

[Figure 12 about here.]

The impact of the introduction of a probation period on the threshold (Figure 13) and hazard (Figure 14) are again qualitatively the same as in the 2-qualities case.

[Figure 13 about here.]

[Figure 14 about here.]

4.3.2. *The impact of the discount factor.* Because the discount factor has essentially the same effects as the separation cost, it should not come as a surprise that the effects of a change in the discount factor on the threshold (Figure 15) and separation hazard (Figure 16) are again qualitatively similar to the ones found in the 2-qualities case.

[Figure 15 about here.]

[Figure 16 about here.]

4.3.3. *The impact of the observation variance.* As shown in Figure 17, an increase in the observation variance reduces the variance of the transition function at low lengths and increases it thereafter. This is because at small lengths, a higher observation variance makes the agent rely more on their length 0 prior belief (which has, by assumption, the same standard deviation of 5, regardless of the observation variance): the higher the observation variance, the more unlikely that any information acquired at the next time step will make the agent deviate much from their prior belief. Because a higher observation variance implies that the agent acquires information at a slower pace (each observation is less informative), the agent's belief at higher tenures is less precise; therefore, it is more likely that the agent should hold a belief with a different mean at the next period, and so the transition function has a larger variance.

If, as predicted, the threshold of separation is barely affected, then we expect that an increase in the observation variance decreases the hazard of separation at short lengths, and, possibly[24], increases it at higher lengths.

[Figure 17 about here.]

When examining the impact of an increase in the variance of the observation on the separation threshold (Figure 18), one observes quasi no effect at very low lengths, and a slight negative effect thereafter.

[24]As already pointed out in section 2.5.2, the effect of a change in the transition function at long lengths cannot be predicted precisely.

Thus, at low lengths, when the observation variance increases, there is no change in the threshold, and the variance of the transition function decreases; therefore the hazard of separation decreases. At higher lengths, the effects on the threshold and the transition function go in opposite directions, making the effect on the separation hazard even more uncertain.

[Figure 18 about here.]

Plotting the separation hazard in Figure 19, we see that the hazard indeed decreases with the observation variance at low lengths, showing that the effect on the transition function dominates the effect on the threshold. At higher lengths the hazard increases with the observation variance. Overall, the qualitative impact on the firing hazard is the same as in the 2-qualities case, despite somewhat different effects on the separation threshold.

[Figure 19 about here.]

4.3.4. *The impact of the prior variance.* An increase in the prior variance increases the variance of the transition function at both high and low lengths (Figure 20), with a vanishing effect at higher lengths. This is very intuitive: since there is a higher match quality variance, all other things equal, agents expect their next period beliefs to vary more. And, since in this model the variance of the transition function converges to 0, variances converge together to 0. We thus expect, all other things equal, a sizable increase in the hazard of separation at low length and, possibly, a smaller positive effect a higher lengths.

[Figure 20 about here.]

The separation threshold decreases with a higher variance of the prior at low lengths, and increases at higher lengths (Figure 21). The effect is however small.

[Figure 21 about here.]

Figure 22 shows that the hazard increases with the variance of the prior at low lengths and that hazards converge to 0 at higher lengths. Thus, again the effect of the transition function dominates the threshold effect at low lengths.

[Figure 22 about here.]

While the model examined in this section is more general than the two match qualities model, the effect of parameters on the hazard of separation is qualitatively very similar in both models. Thus, if one is interested in the effect of parameters such as the separation cost on the hazard of separation, the simpler 2-qualities model can be used with little loss of predictive power.

## 5. Time-varying quality, normally distributed

So far, I have assumed that match quality is constant over time. This is however not very realistic in most settings, for at least two reasons. First, some random shocks could affect the quality of a relationship: for example, tastes and needs (or demand conditions and technology for firms) can change over time in an unpredictable way, and affect match quality. Second, partners typically adapt to each other while in a relationship: they learn how to be more productive in this relationship, and so match quality may systematically improve, at least in the beginning of a relationship.

5.1. **Model specification.** Specifically, match quality is assumed to evolve over time according to the following AR(1) process:

$$q_k = \rho q_{k-1} + c + \epsilon^q_{k-1} \tag{5.1}$$

where $\epsilon^q_k \sim N(0, \sigma_p)$. $c$ is a deterministic drift. Note that, with $\sigma_p = 0$, $\rho = 1$ and $c = 0$, we obtain the model from the previous section; so this model is indeed a generalization. The observation is defined as in

(4.1), except that match quality is now allowed to vary over time. The observation is therefore defined as:

$$z_k = q_k + \epsilon_k^z \tag{5.2}$$

The best estimate $\hat{q}_k$ of $q_k$ given (5.1) and (5.2) is given by the Kalman filter solutions (see Arulampalam et al. (2001)).

$$P(q_k|z_{1:k}) = N(q_k, \hat{q}_k, \sigma_k) \tag{5.3}$$

$$P(q_{k+1}|z_{1:k}) = N(q_{k+1}, \hat{q}_{k+1|k}, \sigma_{k+1|k}) \tag{5.4}$$

where

$$\hat{q}_{k+1|k} = \rho\hat{q}_k + c \tag{5.5}$$

$$\sigma^2_{k+1|k} = \sigma^2_p + \rho^2\sigma^2_k \tag{5.6}$$

$$\hat{q}_{k+1} = \hat{q}_{k+1|k} + K_{k+1}(z_{k+1} - \hat{q}_{k+1|k}) \tag{5.7}$$

$$\sigma^2_{k+1} = (1 - K_{k+1})\sigma^2_{k+1|k} \tag{5.8}$$

In equations (5.7) and (5.8), $K_{k+1}$ is the Kalman gain and is defined as:

$$\begin{aligned} K_{k+1} &= \frac{\sigma^2_{k+1|k}}{\sigma^2_{k+1|k} + \sigma^2_{obs}} \\ &= \frac{\sigma^2_p + \rho^2\sigma^2_k}{\sigma^2_p + \rho^2\sigma^2_k + \sigma^2_{obs}} \end{aligned} \tag{5.9}$$

As previously, because we work in the belief space, we need to express $z_{k+1}$ as a function of other variables. Using equations (5.7) and (5.5):

$$\begin{aligned} z_{k+1} &= \frac{\hat{q}_{k+1} - \hat{q}_{k+1|k}}{K_{k+1}} + \hat{q}_{k+1|k} \\ &= \frac{\hat{q}_{k+1} - (\rho\hat{q}_k + c)}{K_{k+1}} + \rho\hat{q}_k + c \end{aligned} \tag{5.10}$$

The probability of transition from $\hat{q}_k$ to $\hat{q}_{k+1}$, i.e. the belief transition function, is:

$$P(z_{k+1}|\hat{q}_k) = N(\rho\hat{q}_k + c, \sqrt{\rho^2\sigma_k^2 + \sigma_p^2 + \sigma_{obs}^2}) \tag{5.11}$$

Because the computations are performed in the belief space and not the observation space, the expression $\rho^2\sigma_k^2+\sigma_p^2+\sigma_{obs}^2$ is not exactly the variance of the transition function. As previously, the variance of the transition function at length $k$, $VT_k$, can instead be calculated using the definition of the variance of a probability distribution with a known mean:

$$VT_k = \int_{-\infty}^{+\infty} p(z_{k+1}(\hat{q}_{k+1})|\hat{q}_k)(\hat{q}_{k+1} - (\rho\hat{q}_k + c))^2 d\hat{q}_{k+1} \tag{5.12}$$

Note that, due to the assumptions made about the evolution of match quality, $\sigma_k^2$ no longer necessarily decreases with $k$, as it did in the constant match quality case, i.e. beliefs do not necessarily get more precise as length increases. This is because while the agent accumulates observations, match quality changes, and therefore whether the belief gets more precise as relationship length increases depends on whether observations are sufficiently informative given the parameters of the match quality process. More precisely, we have:

$$\sigma_{k+1}^2 < \sigma_k^2 \Leftrightarrow (\sigma_p^2 + \rho^2\sigma_k^2)(\sigma_{obs}^2 - \sigma_k^2) - \sigma_{obs}^2\sigma_k^2 < 0 \tag{5.13}$$

This implies that we can extract conditions over parameters under which the variance of the belief decreases from length $k$ to length $k+1$. Note that the variance of the transition function in equation (5.11) is a function of $\sigma_k$, so that the inequality above has some influence on the variance of the transition function, and therefore on the hazard of separation. Condition (5.13) is however not sufficient to predict how the variance of the transition function evolves with length, because the variance of the transition function also depends directly on $\rho^2$, $\sigma_p^2$ and $\sigma_{obs}^2$, and is given by equation (5.12).

5.2. **Results.** The belief state discretization used here is the same as in section 4 (constant match quality, normally distributed).

The parameters used in the reference case are to be found in Figure 1.

Before examining the results, it is useful to comment on the general properties of the transition function variance when match quality follows an AR(1) process. As mentioned before, the variance of the transition function could either increase or decrease with relationship length. Given the parameters used in the reference case, the variance of the transition function decreases with length but ultimately converges to a stable value, as shown by Figure 29 (in the case where the observation standard deviation is 10). If the variance of the transition function increases with length, and I will show some cases where this happens, the variance of the transition function also converges to a stable value. The relationship between the variance of the transition function and length is explained by the fact that, in the beginning of a relationship, agents learn about match quality, combining their prior with their observations; but the stochastic aspect of the quality process implies that agents eventually hit a wall such that they cannot further improve the precision of their knowledge. Even if the agent's knowledge of today's match quality was perfect, match quality would change tomorrow according to equation (5.1)and so the transition function must always have a non-zero variance. For the hazard of separation, if the threshold does not vary much over time, as is the case here, this implies that the hazard of separation eventually reaches a plateau and no longer decreases. This is an important difference with the constant quality case, where the hazard of separation converges to 0 as length increases.

One more remark is in order concerning the transition function. In the constant quality case, the mean was $\hat{q}$. In this case the mean is $\rho\hat{q}+c$, which in the reference case is $.99\hat{q} < \hat{q}$. This implies that, in the reference case, match quality slightly decreases over time. Together with the fact that the variance of the transition function does not converge to 0, this also accounts for higher hazards of separation at longer lengths in the time-varying match quality model relative to the constant quality models.

5.2.1. *The impact of separation costs.* As in all cases examined so far, higher separation costs decrease the separation threshold (Figure 23) and, thus decrease the separation hazard as well (Figure 24).

[Figure 23 about here.]

[Figure 24 about here.]

The effects of a probation period on both threshold (Figure 25) and hazard (Figure 26) are also qualitatively similar to the effects found in the cases examined before.

[Figure 25 about here.]

[Figure 26 about here.]

5.2.2. *The impact of the discount factor.* Not surprisingly, given that the effects of separation costs are the same as before, the effects of the discount factor are also the same as usual (Figures 27 and 28).

[Figure 27 about here.]

[Figure 28 about here.]

5.2.3. *The impact of the observation variance.* The observation variance has an impact on the transition function (Figure 29) that is similar to the one found in the section 4 constant quality case[25] for short lengths: the variance of the transition function decreases with the variance of the

[25] From now on, when referring to "the constant quality case", I will mean the case examined in section 4.

observation, and this for the same reason as in the constant quality case. As length increases, variances converge to different values, but in such a way that the variance of the transition function remains slightly higher for lower observation variances. This result is thus different from the constant quality case where at high lengths a higher observation variance was associated with a higher variance of the transition function. This different effect comes from the role played by the stochastic aspect of the match quality process. Intuitively, after a while, it does not matter anymore what the belief was at period 0 since match quality has evolved dramatically. Therefore, at high length, the agent's current belief plays the same role the prior played in the constant quality case. This explains why a larger observation variance leads to a lower transition function variance: with a higher observation variance, the observation that the agent will get next period is less likely to change their current belief. Given its impact on the variance of the transition function, a higher observation variance should all other things equal decrease the hazard of separation at low length.

[Figure 29 about here.]

The separation threshold (Figure 30) decreases with observation standard deviation. The overall effect is qualitatively very similar to the effect shown in the constant quality case.

[Figure 30 about here.]

The overall effect of an increase in observation standard deviation on the separation hazard is limited, as shown in Figure 31. Basically, a higher observation standard deviation shifts the whole separation hazard slightly to the right. The fact that the hazard is lower at short lengths for a higher observation variance is not surprising since the transition function variance and the threshold both decrease with observation variance. It is however less obvious why the hazard at higher lengths *increases* with the observation variance, since both the threshold and the observation variance *decrease* with observation variance at *all* lengths. The explanation is given by the consideration of the evolution of the distribution of continuing relationships with length. Intuitively, a noisier observation does not allow the agent to detect the "lemons" as fast and efficiently, which means that the hazard of separation at low lengths is smaller. On the other hand, since with a noisier observation the agent has not been able to sort out the lemons so well in the beginning of relationships, it knows that, once the precision of its knowledge no longer improves, there are more lemons left among the continuing relationships. This is what drives the higher hazard of separation at longer relationship lengths. To understand this in the specific context of the model, let's look at Figure 32. At length 2, we observe that the distribution of continuing relationships with an observation standard deviation of 12 is narrower than the distribution with an observation standard deviation of 8[26]. This implies a lower separation hazard for "12" than for "8", and can be seen by the fact that the part cut to the left of the distribution (which is exactly equal to the hazard) is fatter with an observation standard deviation of 8. This also implies that fewer low quality relationships are terminated at length 2 with "12" versus "8". At length 10, the "12" distribution of continuing relationships has a higher density in the neighborhood of the threshold, implying that the "12" hazard is now *bigger* than the "8" hazard. The higher proportion of low quality relationships at length 10 in the "12" case is explained by the fact that overall fewer low quality relationships have been terminated before length 10, and so there are more low quality relationships to terminate at length 10 and later[27].

[26]This is because, as shown in Figure 29, the variance of the transition function decreases with the observation variance.

[27]Notice that, since the prior distribution is held constant, the distribution of relationships at short lengths is not affected by the change in the observation variance. In other terms, in cases "8" and "12", the agent starts with an identical density of actual matches below the threshold, so it makes sense to say that fewer of the low quality matches have been terminated in the "12" case.

Besides, because the agent's belief does not get more precise after length 10, this difference in the proportion of low quality relationships between the "8" and the "12" cases at length 10 will tend to persist over time.

[Figure 31 about here.]

[Figure 32 about here.]

5.2.4. *The impact of the prior variance.* As shown in Figure 33, a higher prior variance increases the variance of the transition function at all lengths, as in the constant quality case. However, the variance here converges to the same value, for all the prior variance considered. The convergence is explained by the fact that, as already mentioned, if match quality is time-varying, the prior ceases to matter after a while. Note that with a prior standard deviation smaller than 5, the variance of the transition function increases with length.

[Figure 33 about here.]

Moreover, as in the constant quality case, a higher standard deviation for the prior decreases the threshold for separation at very short lengths and increases it thereafter (Figure 34).

[Figure 34 about here.]

With a higher standard deviation of the prior, the hazard of separation is higher at low length, and the maximum of the hazard occurs earlier (Figure 35). The effect on the hazard is slightly different from the effect observed in the constant quality case, in as much as, at higher lengths, the hazard is higher and not lower for a lower standard deviation of the prior. This is explained by the fact — mentioned earlier when discussing the impact of the observation variance — that a lower variance of the transition function at short lengths leads to a permanently higher separation hazard all other things equal.

[Figure 35 about here.]

5.2.5. *The impact of the drift.* The drift in the match quality process does not affect the variance of the transition function. It does however change its mean in a straightforward additive way, as seen in equation (5.11).

A larger drift has a positive effect on the separation threshold, effect which is roughly similar for all lengths (Figure 36).

[Figure 36 about here.]

A larger drift has a negative impact on the hazard of separation (Figure 37). This is because, for a given threshold, a lower mean for the transition function means that at each length more relationships cross the threshold. Even though the threshold decreases with a smaller drift, this is not enough to countervail the effect on the transition function. Intuitively, the more matches worsen over time, the higher the separation hazard.

[Figure 37 about here.]

5.2.6. *The impact of the process variance.* As shown in Figure 38, an increase in the standard error of the process leads to a higher variance of the transition function at all lengths. Moreover, when the process variance is below 3, the variance of the transition function decreases with length, whereas it increases with length if the process variance is greater than 3. Variances converge to a constant value, and this value is higher with a higher process variance. All other things equal, this should lead to a smaller hazard of separation at short lengths when the process variance is smaller.

[Figure 38 about here.]

A higher standard deviation of the process decreases the threshold (Figure 39), and this effect is less strong at very short lengths.

[Figure 39 about here.]

Overall, the impact of the process variance on the separation hazard is positive at short lengths and negative thereafter (Figure 40). The impact of an increase in process variance on the hazard is mainly explained by the fact that a higher process variance increases the variance of the transition function. The mechanism at play is identical to the one examined for a *decrease* in the observation variance, which also leads to an increase in the variance of the transition function (see section 5.2.3): namely, a higher variance of the transition function implies that more low quality relationships are dissolved at low lengths, and, precisely because of this, fewer relationships are left to be dissolved at higher length. Finally, one may wonder why the hazard of separation even increases for a process variance of 1: this is because, as $\rho = .99$, match quality deterministically decreases over time, so if the process variance is close to 0, then this decrease in quality drives the increase in the separation hazard.

[Figure 40 about here.]

5.2.7. *The impact of the AR(1) parameter of the process.* The AR(1) parameter $\rho$ of the process has a positive impact on the variance of the transition function, as illustrated by Figure 41. This positive impact is however small at low lengths and increases thereafter. Moreover, a bigger AR(1) parameter increases the mean of match quality given previous match quality.

[Figure 41 about here.]

A higher $\rho$ increases the separation threshold, as seen in Figure 42. This effect is roughly equal for all relationship lengths.

[Figure 42 about here.]

Finally, a higher $\rho$ has little effect on the hazard at very low lengths, but decreases it for longer lengths (Figure 43). This is because the effect of $\rho$ on mean match quality given previous match quality (see equation 5.1) dominates: a smaller $\rho$ implies that relationship quality deteriorates faster, and thus increases the separation hazard.

[Figure 43 about here.]

While the effects of changes in separation costs or the discount factor on the hazard of separation are qualitatively similar in this model compared to the others, the effect of parameters entering the transition function are different. Essentially, this is because of the interactions between the extra parameters in this model (i.e. $\rho$, $\sigma_p$ and $c$) and the other parameters entering the transition function.

Therefore, one must ask if this model really yields any practical benefits. One important way in which this model is superior to the ones used in previous sections is that it allows for the hazard of separation not to decline to 0 as relationship length increases. Indeed, such decline to 0 is typically not observed empirically. The substantial reason behind this is that, realistically, match quality is not constant over time, and therefore it is important for a model to account for such time variation.

## 6. Discussion

**6.1. Effort and labor supply.** The models presented have not explicitly integrated agents' efforts. This is a very important issue as match quality could be in part determined by agents' efforts. For example, in the employment relationship, the employee can affect output by supplying more or less unobserved effort, as in Holmstrom(1999). In a formal framework very similar to the one I use here, the latter article shows that labor supply will decline to 0 if an employment relationship continues indefinitely and the worker's ability is fixed. On the other hand, if ability evolves stochastically, labor supply will be positive and stable over time (after an initial period of adjustment). Homlstrom's results imply that the models developed in sections 3 (two constant match qualities) and 4 (normally distributed, constant match quality) are inconsistent in the presence of a serious moral hazard problem; these models assume indeed that the agent (wrongly) believes that the benefits from the relationship do not depend on the partner's effort. If match quality evolves over time as in section 5, then the model is not necessarily inconsistent, even in the presence of moral hazard. Further exploration of this issue is left, however, to future work.

6.2. **General equilibrium.** The analysis developed in this paper is in partial equilibrium; it does not attempt to model the influence of the behavior of each agents on the others. If relationship quality is entirely match specific, then this is not a problem as the prior distribution of match qualities faced by the agent is not influenced by the behavior of other agents. If, however, match quality is at least in part due to some general characteristics that make a partner desirable to all agents, then a change in behavior by other agents is likely to change the distribution of prior match qualities. For example, if firms face higher firing costs and, as a result, decrease their threshold for separation, then the distribution of prior match qualities should have a slightly lower mean since now workers who were terminated and are looking for a new job are a bit worse on average[28]. The feedback mechanism from agents' optimal behavior to the distribution of prior match qualities could in principle be modeled within the framework used here, and it would be useful to do so in future work[29].

Another related issue is that this model does not allow agents not to be in a relationship at all. By assumption, the agent can only continue the current relationship or separate and start a new one immediately. This is an important limitation in contexts such as the labor market where vacancies

[28]This effect is smaller the more workers with no prior experience enter the labor market, and the more match-specific productivity is.

[29]One important challenge is that the feedback from behavior to the distribution of prior match qualities would likely make the latter distribution non normal. Computations are greatly eased if one assumes normality of the distribution of prior match qualities, and to preserve these desirable properties, one would have to devise a meaningful way to approximate the non normal distribution of prior match qualities by a normal one.

and permanent layoffs do exist and are essential for understanding the dynamics of the labor market. The model, however, already contains the tools to analyze these issues, at least in a limited sense. Indeed, one can assume that at length 1 the separation cost is extremely low, and call period 1 the screening period: thus, in period 1 the agent meets a partner, gets a signal about match quality and decides to pursue the relationship or not. This application will be developed in future work and can allow to determine, for example, if firing costs reduce hiring (where hiring means not firing at length 1) more than firing (at lengths 2 and above) and under which conditions this is the case.

6.3. **Learning about match quality, learning on the job, and random shocks.** The model presented in section 5, with time-varying match quality, can simultaneously account for learning about match quality, learning on the job, and random shocks to match quality. These three elements are typically included in separate models in the literature about match quality in the employment relationship. Learning about match quality is thus the main component in Jovanovic(1979). Teulings and van der Ende(2001) develop a model where match quality is subject to random shocks. Using a model that integrates all these empirically relevant effects at the same time[30] is more efficient for empirical analysis because parameters can be determined jointly from a single statistical model. From an empirical perspective, it will typically be difficult to disentangle the effects of these different elements, because of insufficient information about parameters. However, having an integrated model is a useful first step towards finding empirical settings where the effect of a parameter or a set of parameters can be identified.

[30]Nagypal(2004) offers a somewhat different way of integrating these effects in her model.

## 7. Conclusion

This paper has developed a model of optimal matching and separation, allowing for partially observed and time-varying match quality. Despite the limitations discussed in section 6 — some of which could be overcome in future work — the model is already very general and sheds useful light on the mechanisms at play in relationship evolution and dissolution. Specifically, I have shown that, in all models considered, higher separation costs and a lower discount factor decrease the separation threshold and thus the separation hazard. The effect of parameters entering the belief on the separation hazard only depends on the effect of these parameters on the belief transition function[31], i.e. the probability of the agent's holding a certain belief next period given the agent's current belief. In all cases, an increase in the variance of the belief transition function leads to a higher hazard of separation at short lengths. A lower observation variance, a higher variance of the prior, and a higher variance of the error in the AR(1) process all increase the variance of the transition function at short lengths, and thus increase the hazard at short lengths. The effect of parameters entering the transition function at longer lengths is not so clear cut. If match quality is assumed to be constant over time, then the separation hazard converges to 0, and so there will be little effect at higher lengths. If however match quality follows and AR(1) process, then an increase in the variance of the transition function at all lengths typically *lowers* the separation hazard at higher lengths.

The class of models developed here lends itself to applications in various contexts. As already mentioned, domains of choice would be the employment relationship, marriage, and firm-suppliers relationships. Empirically, hazards of separation from an employment relationship and hazards of divorce both increase and decrease over the length of the relationship, but

[31]This is only true if we assume that the expected value of the prior does not change.

do not decline to 0. This implies that, very likely, the underlying match quality is time-varying and separation costs are positive. In general, it is possible to determine which parameters best fit[32] an empirically observed separation hazard and thus gain useful information about the matching process. The model is also useful in predicting the impact of a parameter change on the hazard of separation. For example, in Marinescu(2006), I examined the impact of a change in the probationary period on the hazard of an employment relationship being terminated by the employer.

This paper developed a formal framework with empirical applications in view. It is useful both as a conceptual tool to understand the issues involved in this class of problems in a real world environment, and as a statistical tool for structural estimation. The model will thus hopefully be a starting point for fruitful future empirical work.

[32]We have to keep in mind, however, that the greater the number of unobserved parameters, the less precise the estimates. For example, I already pointed out that one cannot typically distinguish separation costs from the discount factor just by looking at the separation hazard.

## References


[1] Arulampalam S., Maskell S.,Gordon N., Clapp T. (2002), "A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking", *IEEE Transactions on signal processing*, Vol. 50, No. 2, Feb. 2002.

[2] Cassandra A. R. (1998), "Exact and Approximate Algorithms for Partially Observable Markov Decision Processes", Ph.D. thesis, Brown University, May 1998.

[3] Farber H. S. (1994), "The Analysis of Interfirm Worker Mobility", *Journal of Labor Economics*, Vol 12., 554-593.

[4] Hauskrecht M. (2000), "Value-Function Approximations for Partially Observable Markov Decision Processes", *Journal of Artificial Intelligence Research*, 13 (2000) 33-94.

[5] Holmstrom B. (1999), "Managerial Incentive Problems - A Dynamic Perspective", *Review of Economic Studies*, Vol 66 (1), January 1999.

[6] Jovanovic B. (1979), "Job Matching and the Theory of Turnover", *The Journal of Political Economy*, Vol 87., 972-990.

[7] Marinescu I. E.(2006a), "Shortening the Tenure Clock: The Impact of Strengthened U.K. Job Security Legislation", working paper, 2006.

[8] Nagypal E. (2004), "Learning-by-Doing Versus Learning About Match Quality: Can We Tell Them Apart?", working paper, September 2004.

[9] Svarer M. (2004), "Is Your Love in Vain? Another Look at Premarital Cohabitation and Divorce", *Journal of Human Resources* 39(2), 523-536.

[10] Teulings C. N., Van Der Ende, M. A. (2001), "A Structural Model of Tenure and Specific Investments", CESifo Working Paper No. 532, August 2001.

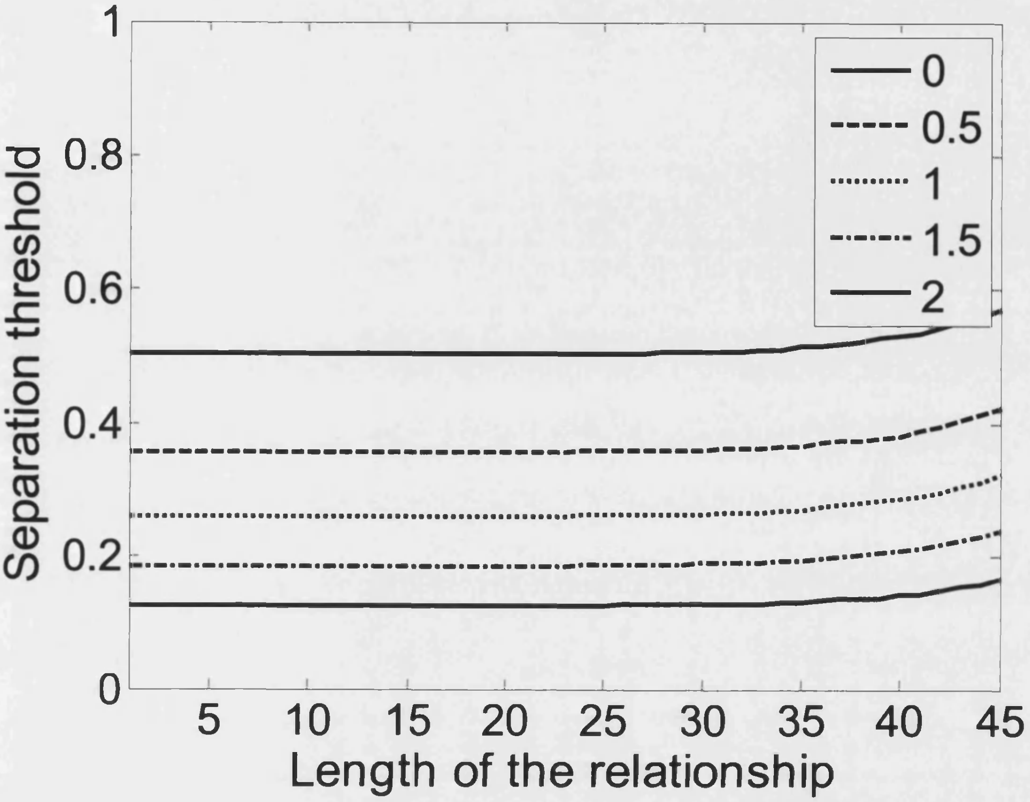
[11] Weiss Y., Willis R. J. (1997), "Match Quality, New Information, and Marital Dissolution", *Journal of Labor Economics*, Vol. 15, No. 1, Part 2: Essays in Honor of Yoram Ben-Porath. (Jan., 1997), pp. S293-S329.

FIGURE 1. Parameters in the reference case

| | 2 match qualities | Normally distributed match qualities | Normally distributed match qualities, with AR(1) process for match quality |
|---|---|---|---|
| Parameters of interest | | | |
| Mean of prior | 0.5 | 30 | 30 |
| Standard deviation of prior | N/A | 5 | 5 |
| Standard deviation of process | N/A | N/A | 2 |
| Dirft of process | N/A | N/A | 0 |
| Auto-correlation of process | N/A | N/A | 0.99 |
| Standard deviation of observation | 4 | 10 | 10 |
| Separation cost | 1.5 | 30 | 30 |
| Discount factor | 0.85 | 0.85 | 0.85 |
| Technical parameters | | | |
| Range of match qualities | [0,1] | [0,60] | [0,60] |
| Number of match quality values | 1001 | 801 | 801 |
| Maximal length | 50 | 50 | 50 |

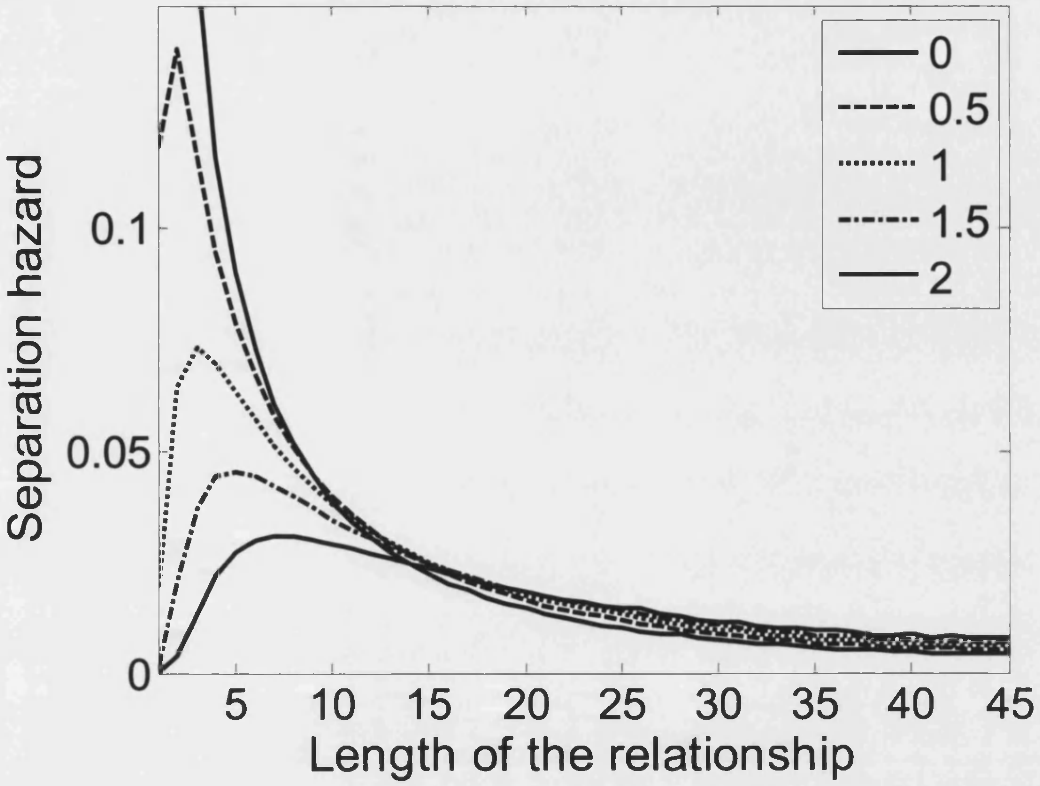
Notes: In the 2 match qualities case, the values of match qualities are defined by $normcdf(X, 0.5, 0.1)$, where $X$ is a vector of 1001 equally spaced values between 0 and 1, and $normcdf(X, 0.5, 0.1)$ is the cumulative distribution function, evaluated at point $X$, of a normal probability distribution with mean 0.5 and standard deviation 0.1. In the other cases, values of match quality are equally spaced over the range.

FIGURE 2. Separation threshold for different separation costs, 2 match qualities



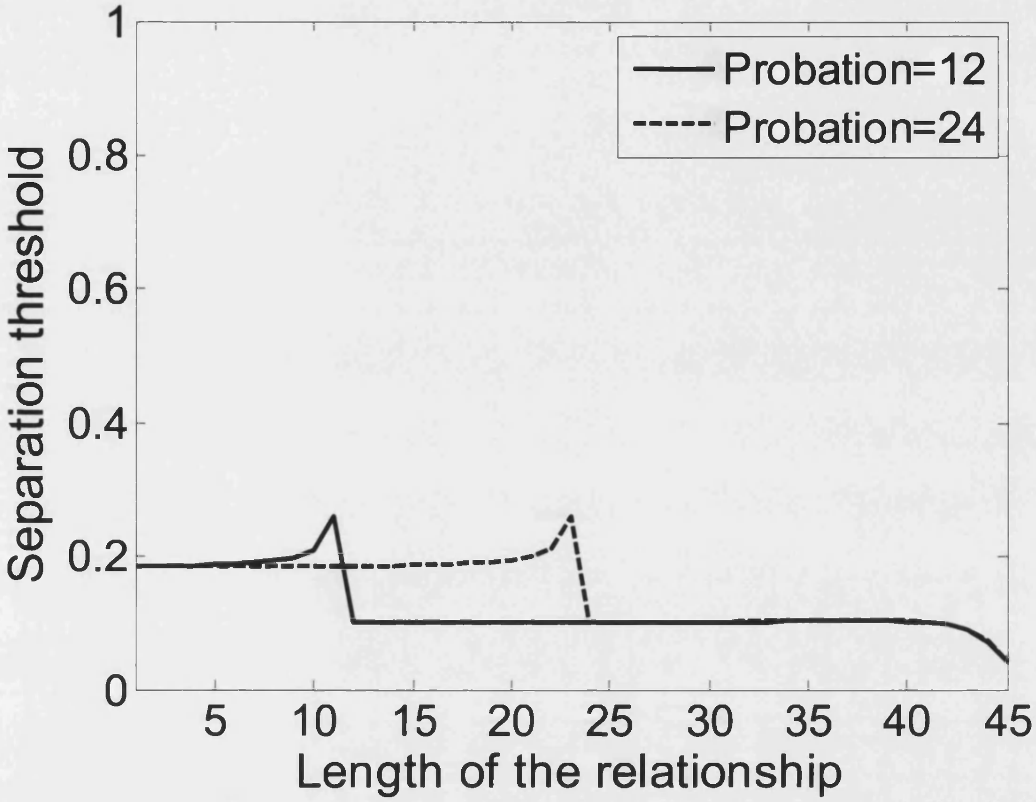

Notes: The figure shows the threshold for different separation costs, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 3. Separation hazard for different separation costs, 2 match qualities



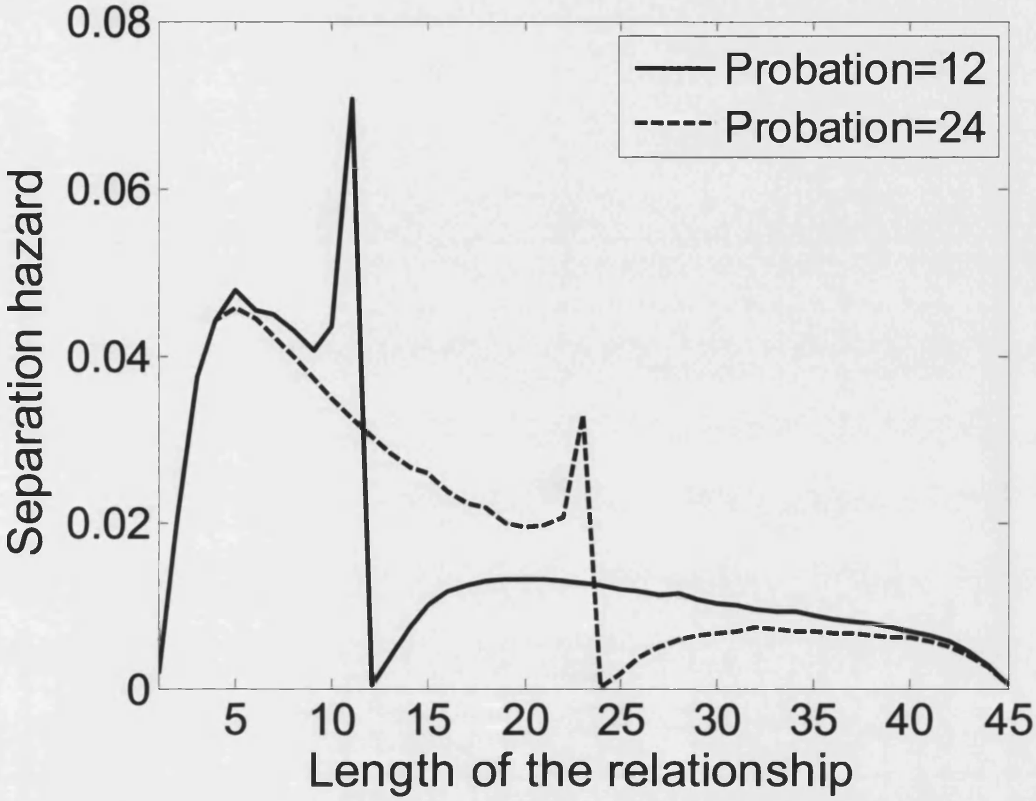

Notes: The figure shows the hazard for different separation costs, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 4. Separation threshold with a probation period, 2 match qualities



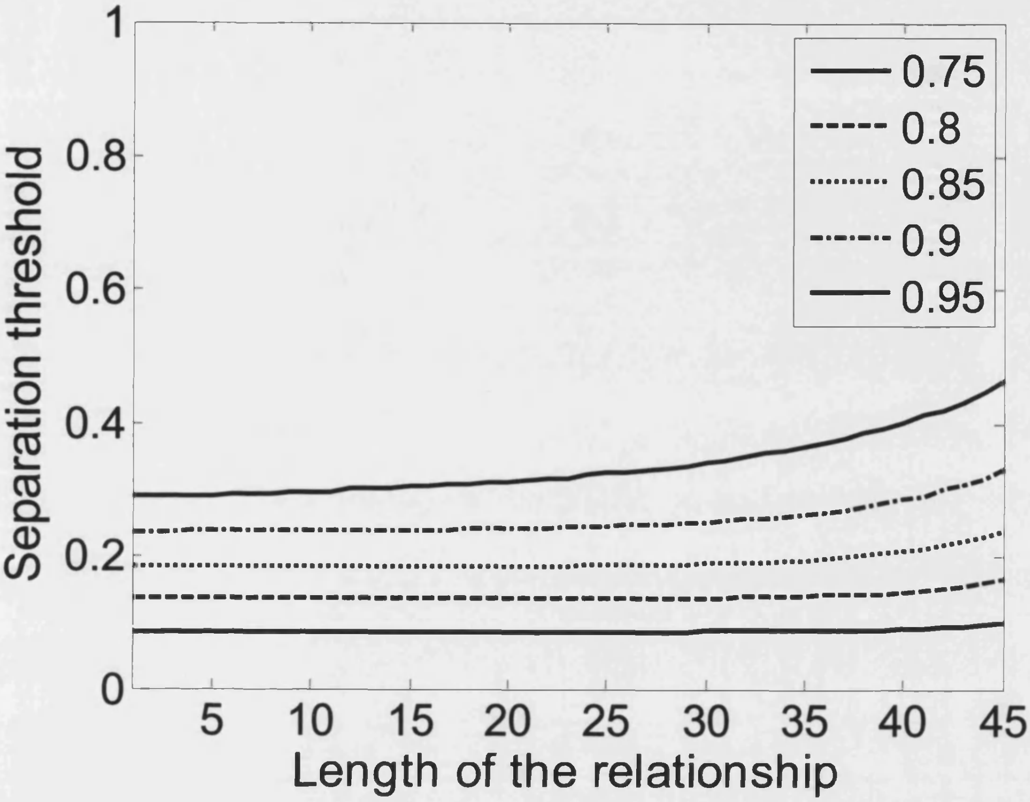

Notes: The figure shows the threshold for different lengths of the probation period, with separation costs being 1.5 during the probation period and 2.3 thereafter. All other parameters are fixed to their reference values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 5. Separation hazard with a probation period, 2 match qualities



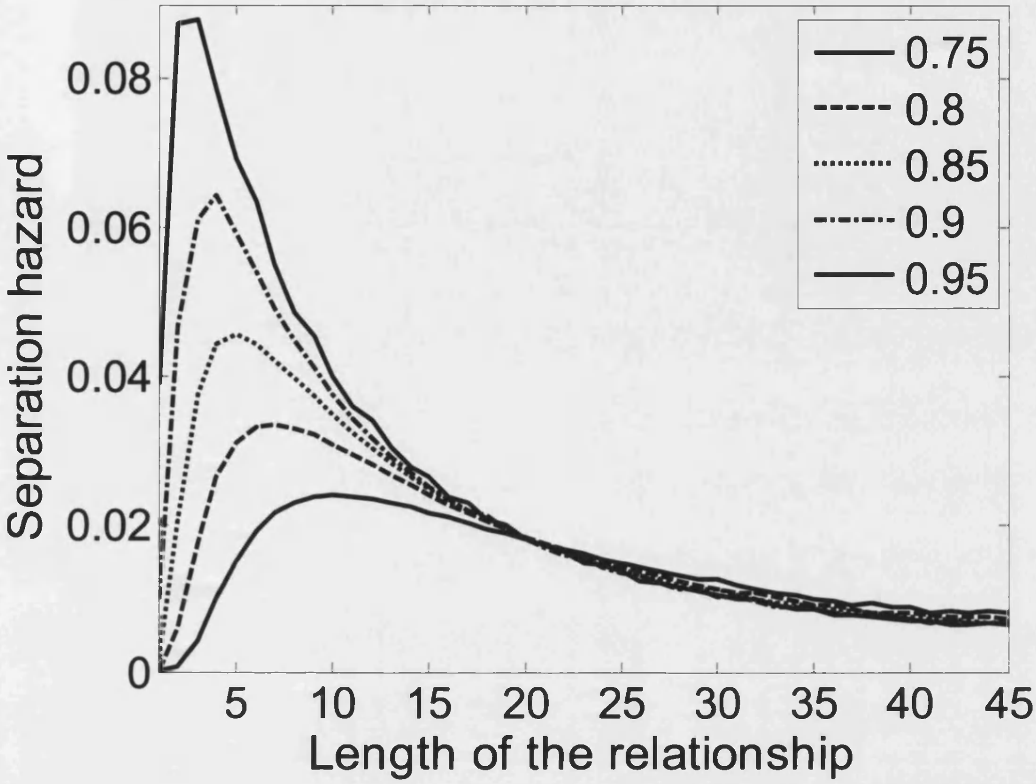

Notes: The figure shows the hazard for different lengths of the probation period, with separation costs being 1.5 during the probation period and 2.3 thereafter. All other parameters are fixed to their reference values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 6. Separation threshold for different discount factors, 2 match qualities



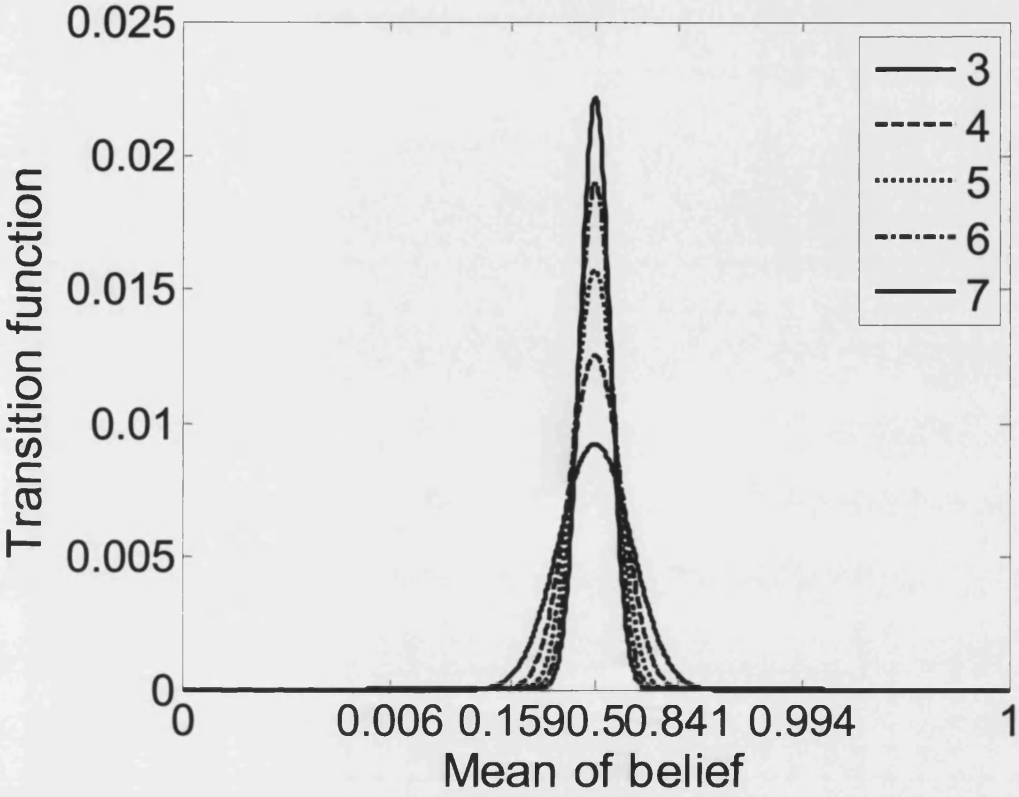

Notes: The figure shows the threshold for different discount factors, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 7. Separation hazard for different discount factors, 2 match qualities



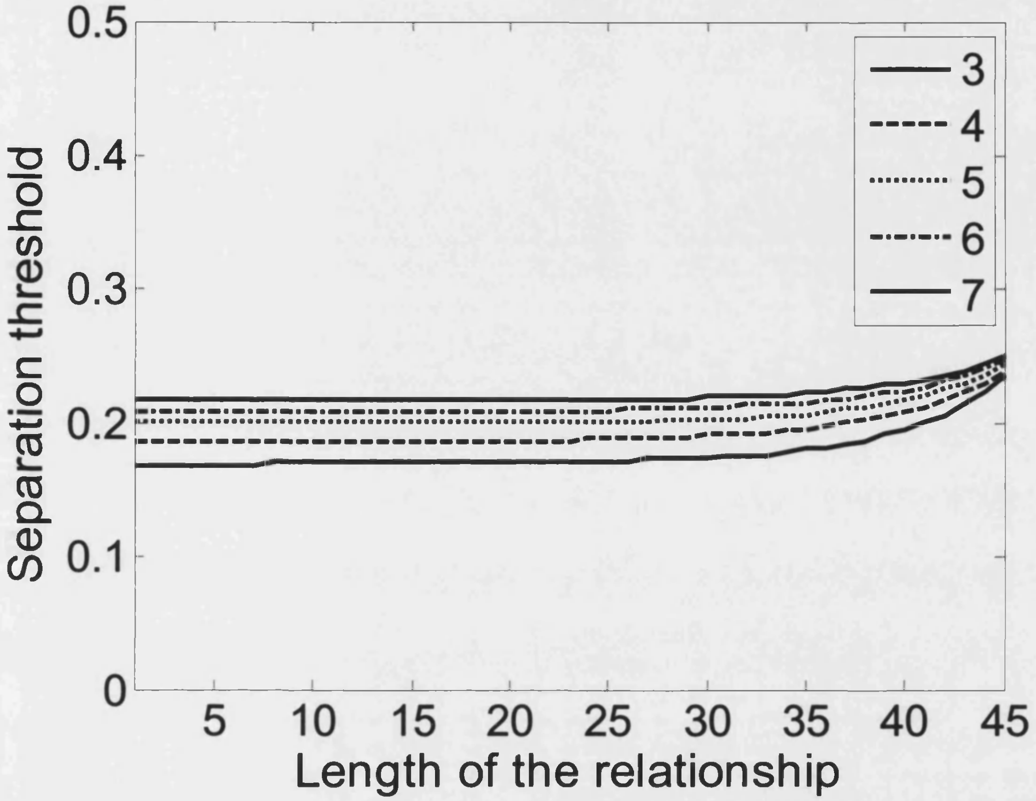

Notes: The figure shows the hazard for different discount factors, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 8. Transition function for different observation standard deviations, 2 match qualities



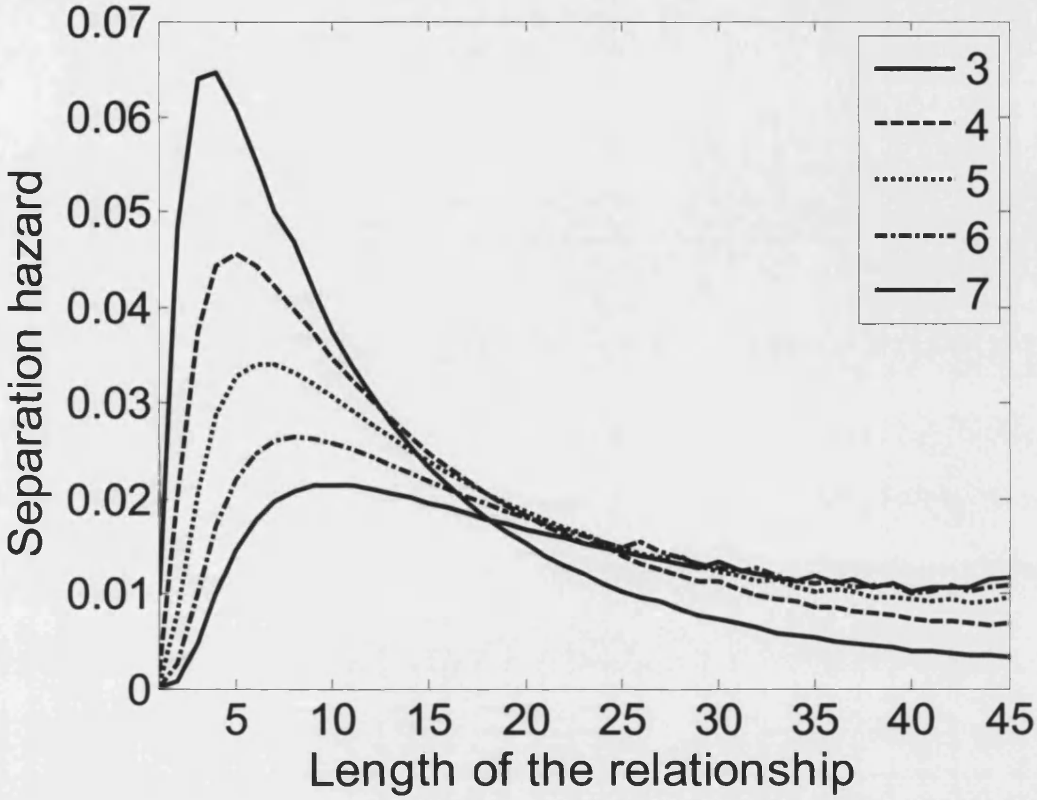

Notes: The figure shows the transition function for different observation standard deviations, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 9. Separation threshold for different observation standard deviations, 2 match qualities



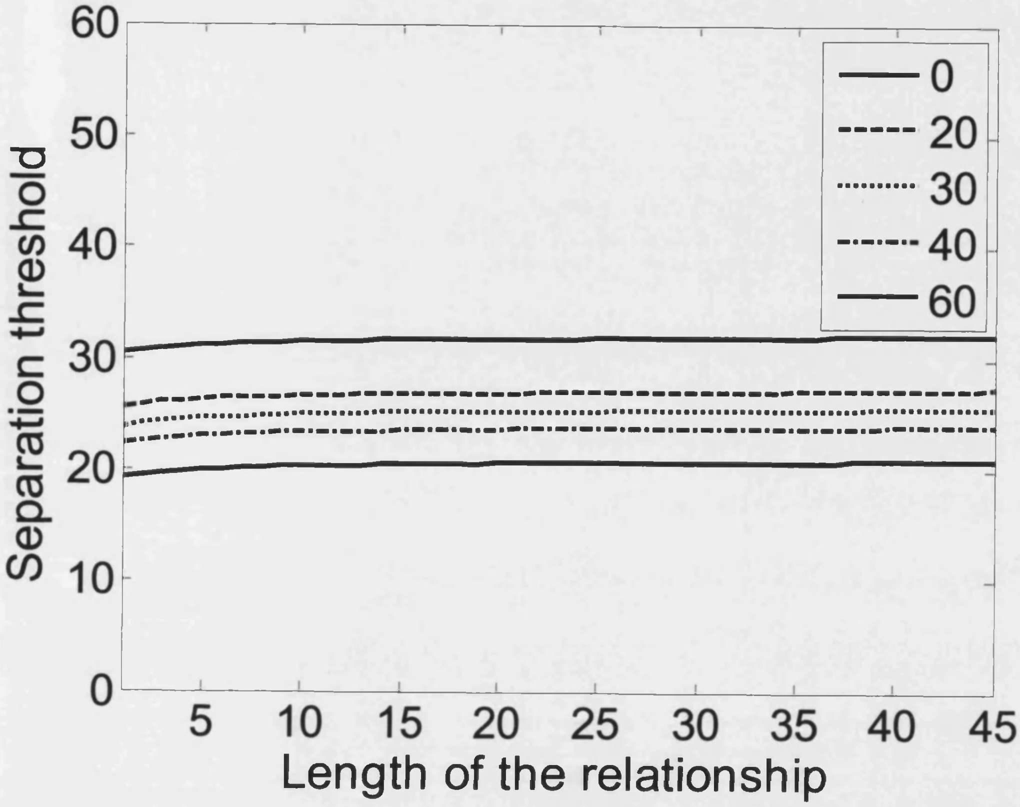

Notes: The figure shows the separation threshold for different observation standard deviations, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 10. Separation hazard for different observation standard deviations, 2 match qualities



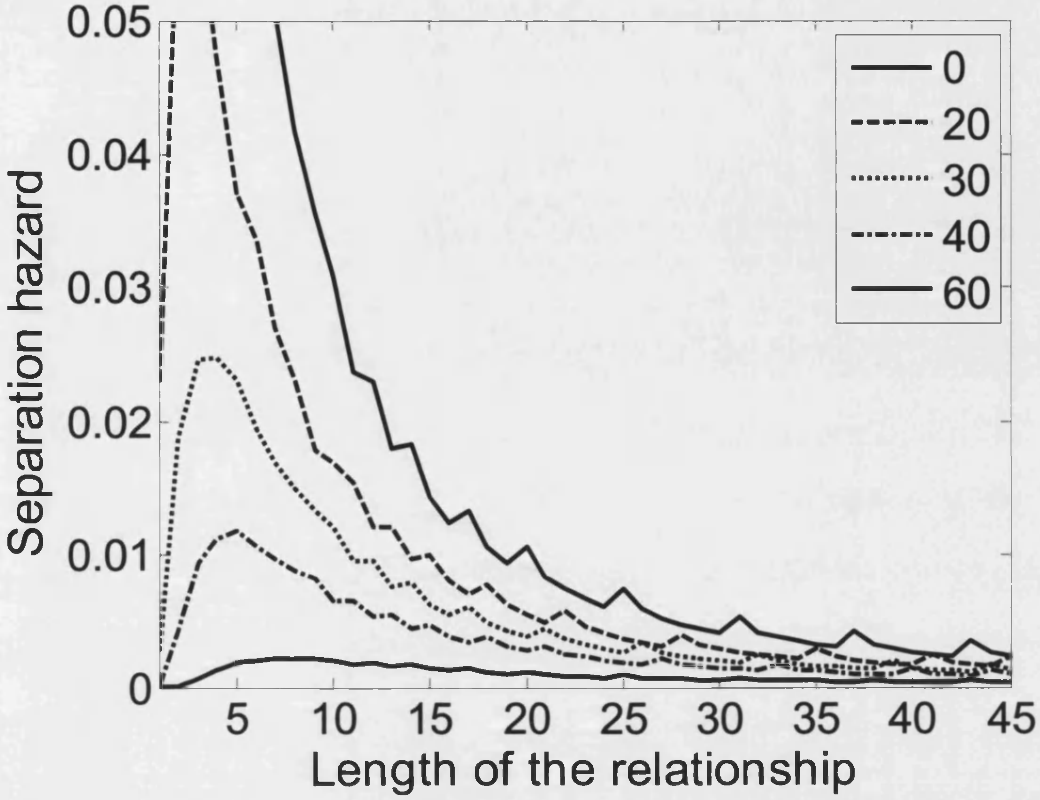

Notes: The figure shows the separation hazard for different observation standard deviations, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 11. Separation threshold for different separation costs, normally distributed match qualities



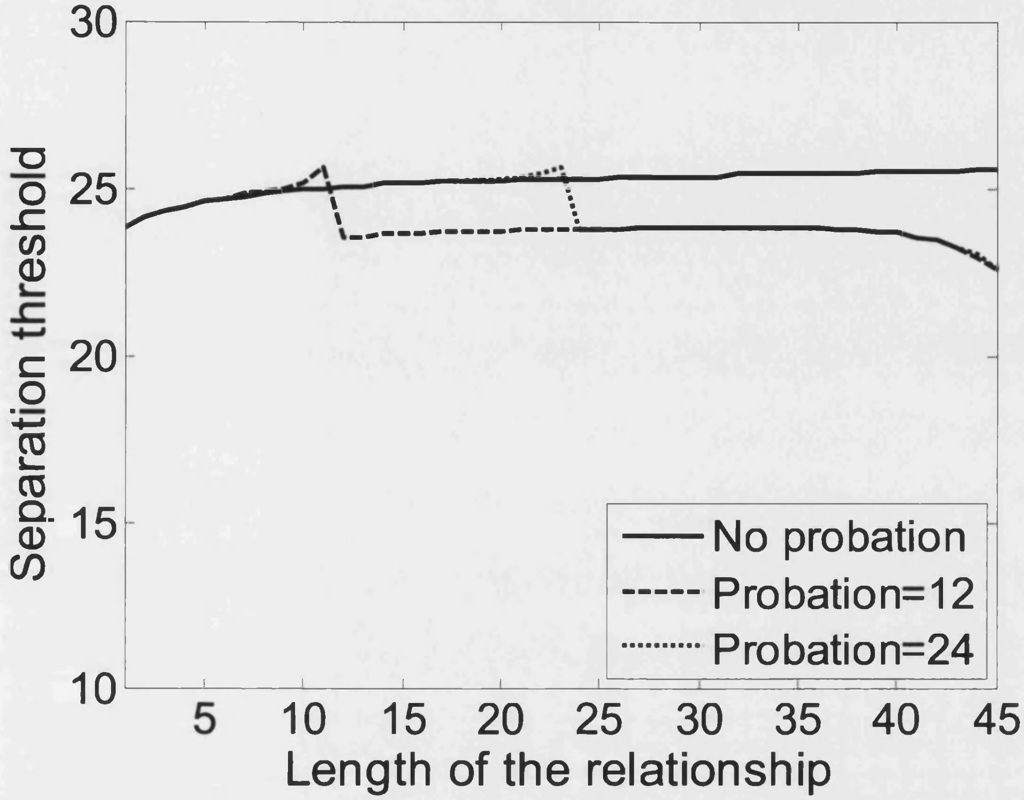

Notes: The figure shows the threshold for different separation costs, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 12. Separation hazard for different separation costs, normally distributed match qualities



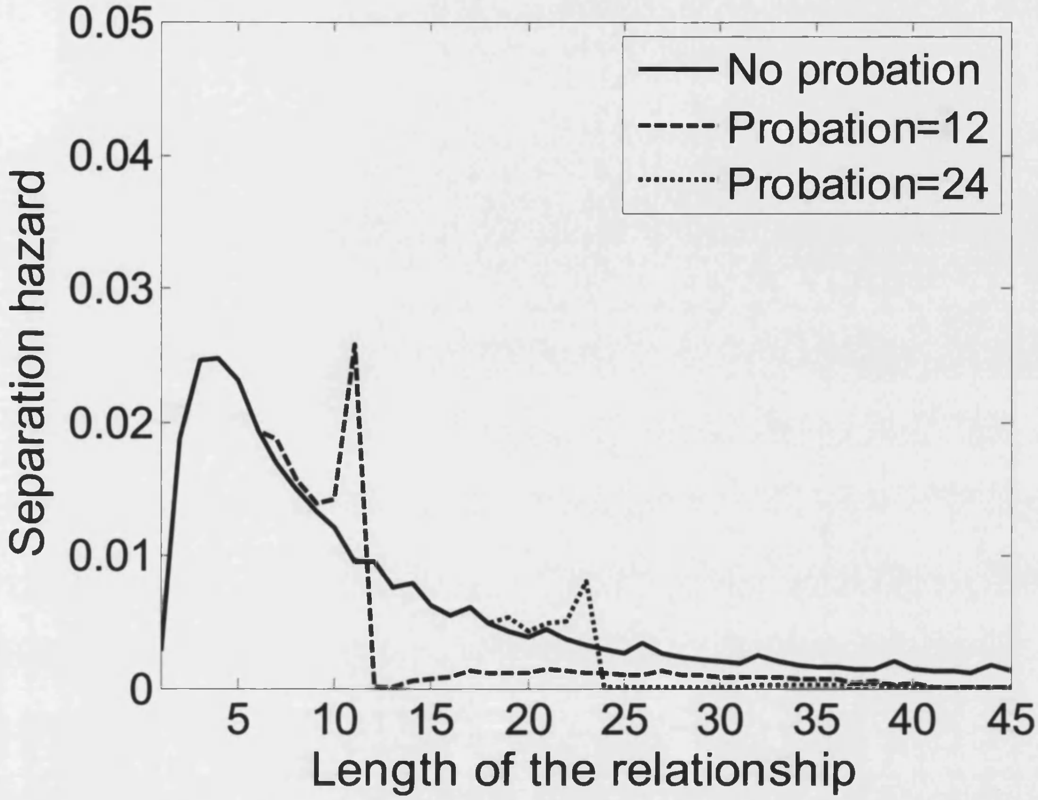

Notes: The figure shows the hazard for different separation costs, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 13. Separation threshold with a probation period, normally distributed match qualities



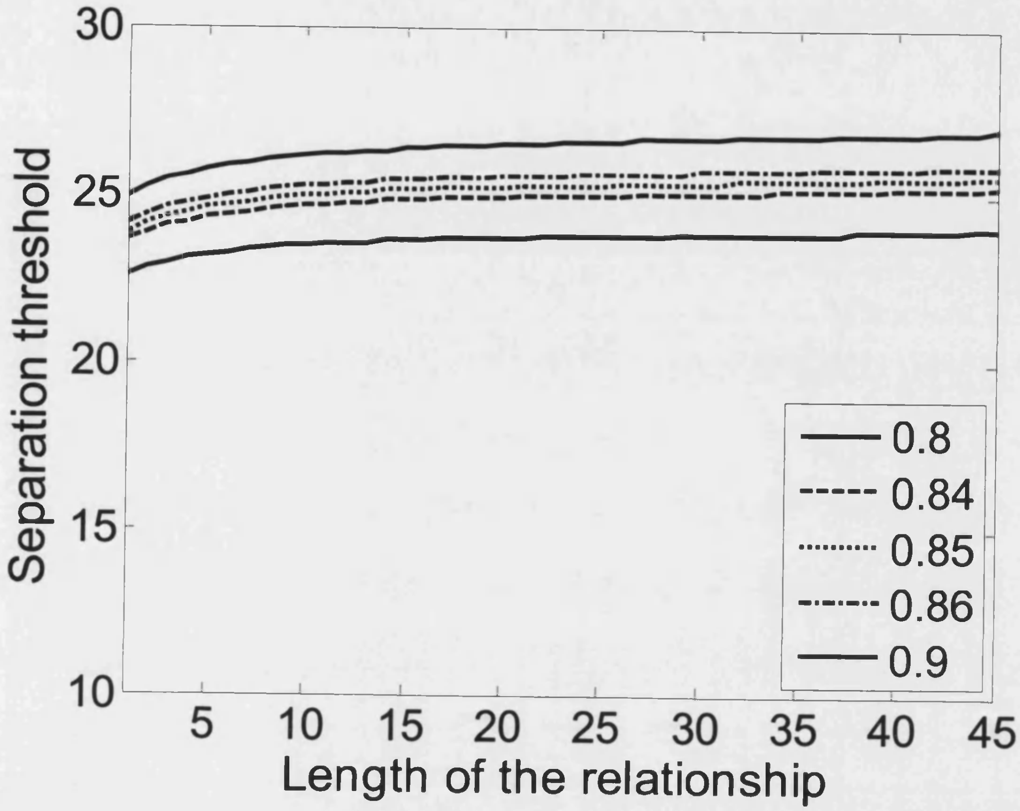

Notes: The figure shows the threshold for different lengths of the probation period, with separation costs being 1.5 during the probation period and 2.3 thereafter. All other parameters are fixed to their reference values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 14. Separation hazard with a probation period, normally distributed match qualities



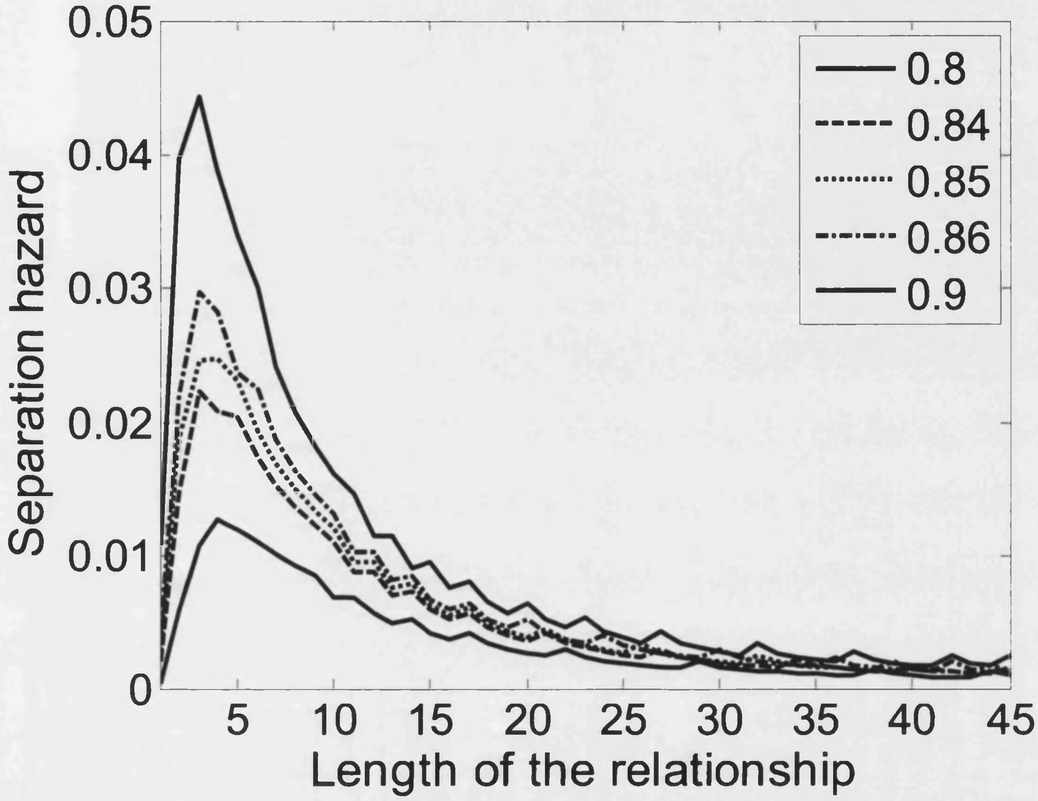

Notes: The figure shows the hazard for different lengths of the probation period, with separation costs being 1.5 during the probation period and 2.3 thereafter. All other parameters are fixed to their reference values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 15. Separation threshold for different discount factors, normally distributed match qualities



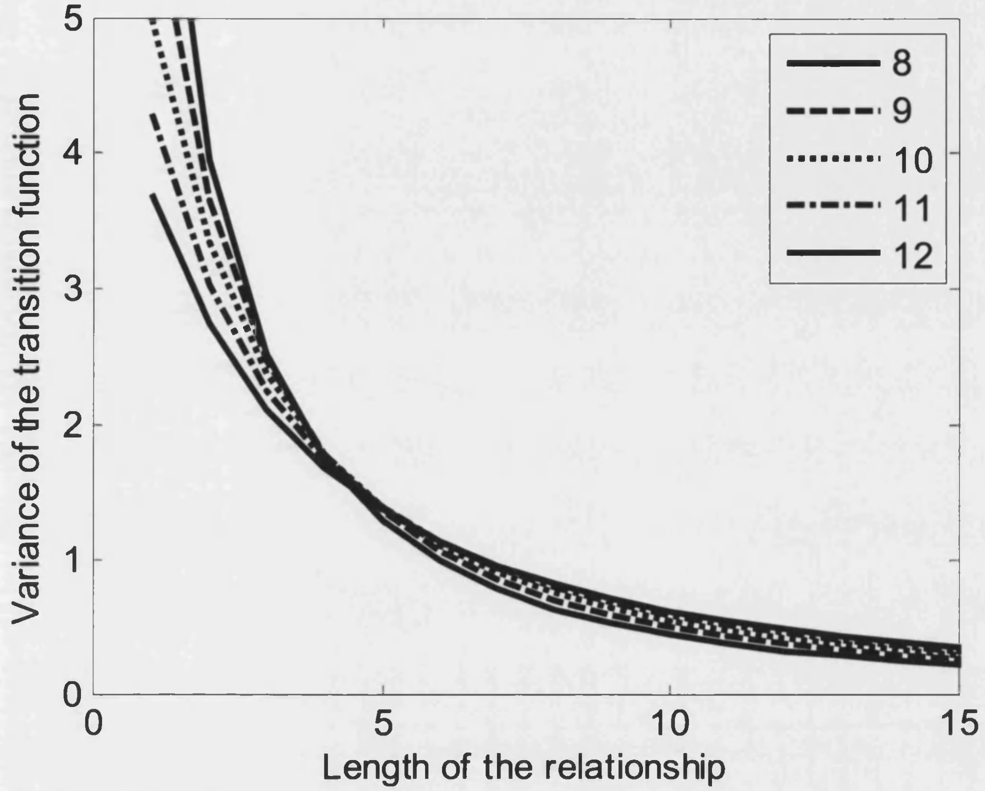

Notes: The figure shows the threshold for different discount factors, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 16. Separation hazard for different discount factors, normally distributed match qualities



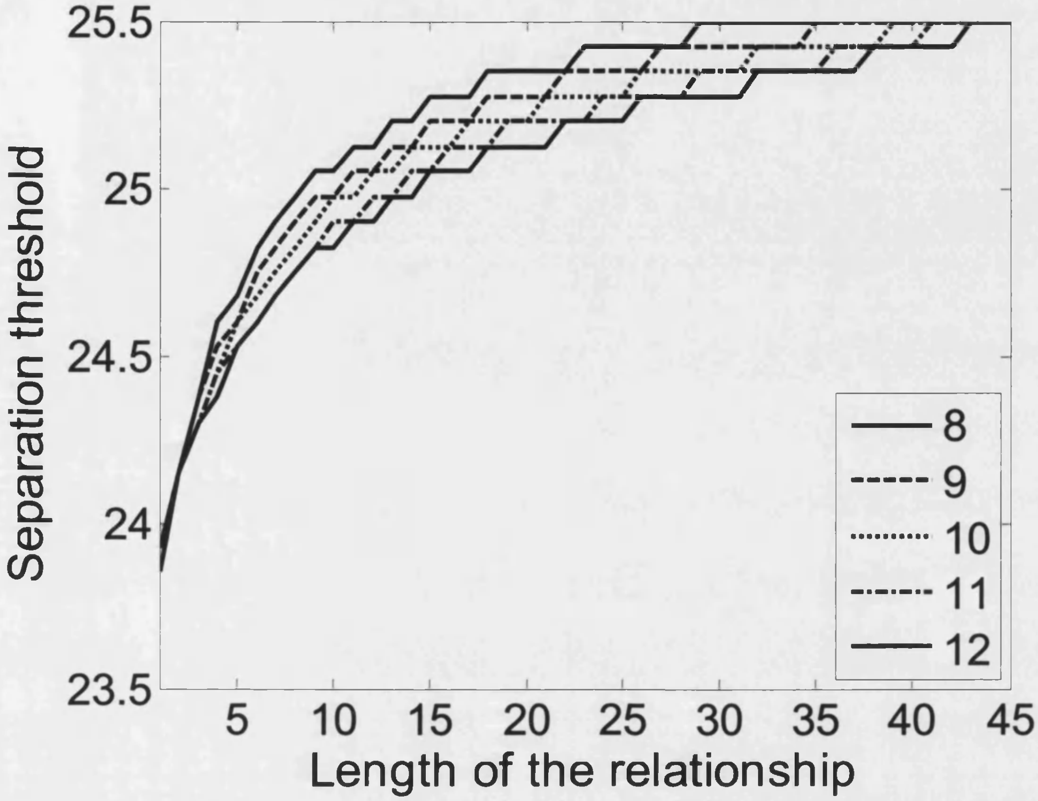

Notes: The figure shows the hazard for different discount factors, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 17. Variance of the transition function for different observation standard deviations, normally distributed match qualities



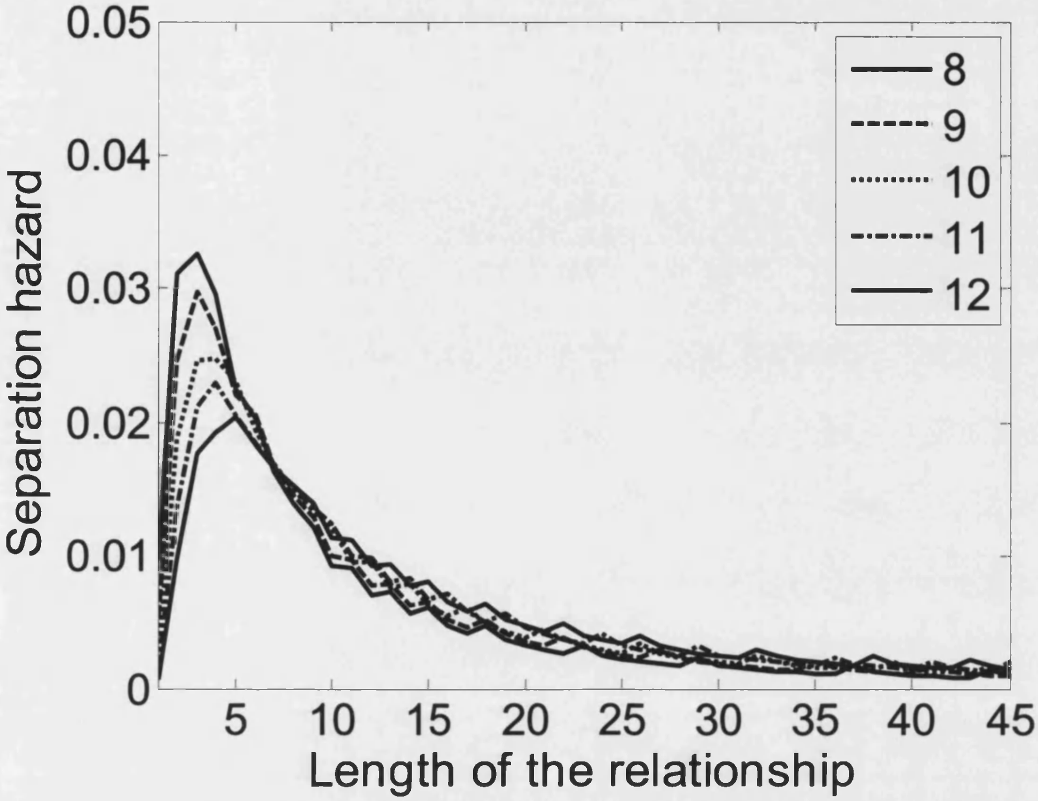

Notes: The figure shows the variance transition function for different observation standard deviations, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 18. Separation threshold for different observation standard deviations, normally distributed match qualities



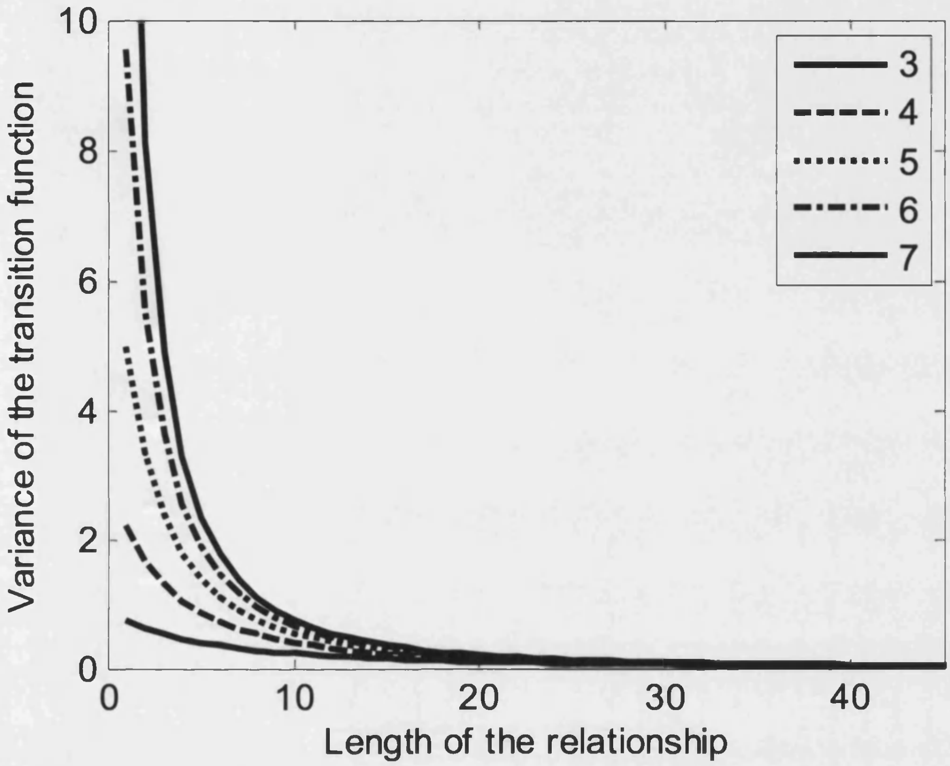

Notes: The figure shows the threshold for different observation standard deviations, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 19. Separation hazard for different observation standard deviations, normally distributed match qualities



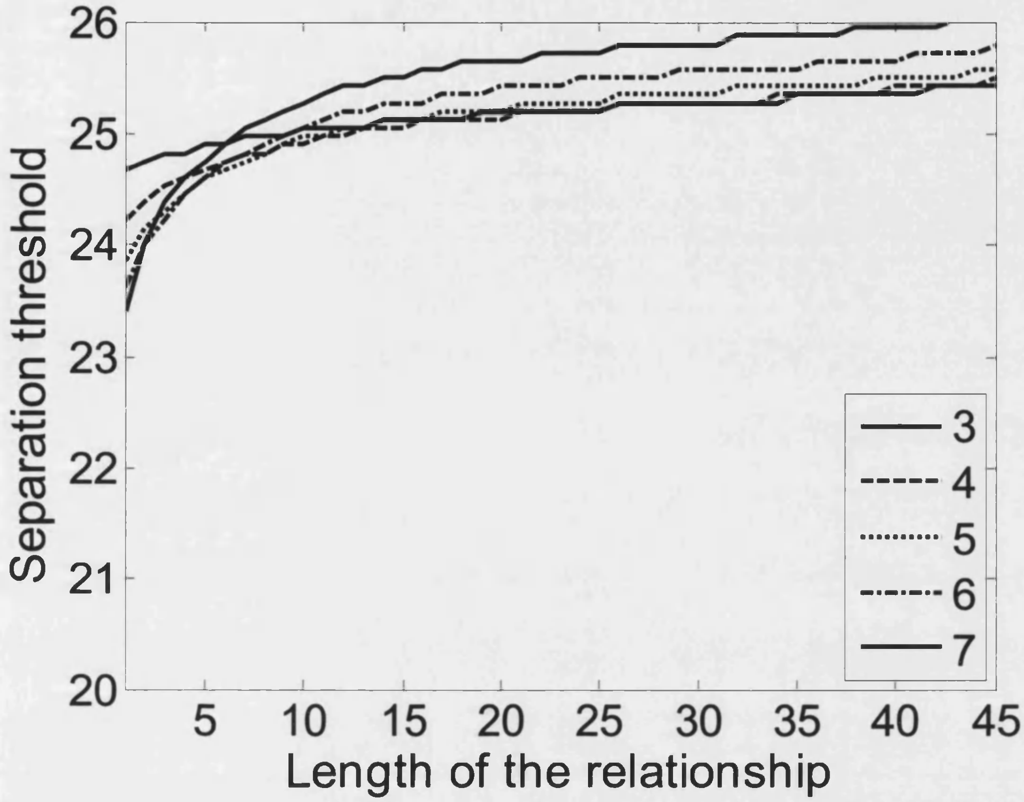

Notes: The figure shows the hazard for different observation standard deviations, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 20. Variance of the transition function for different prior standard deviations, normally distributed match qualities



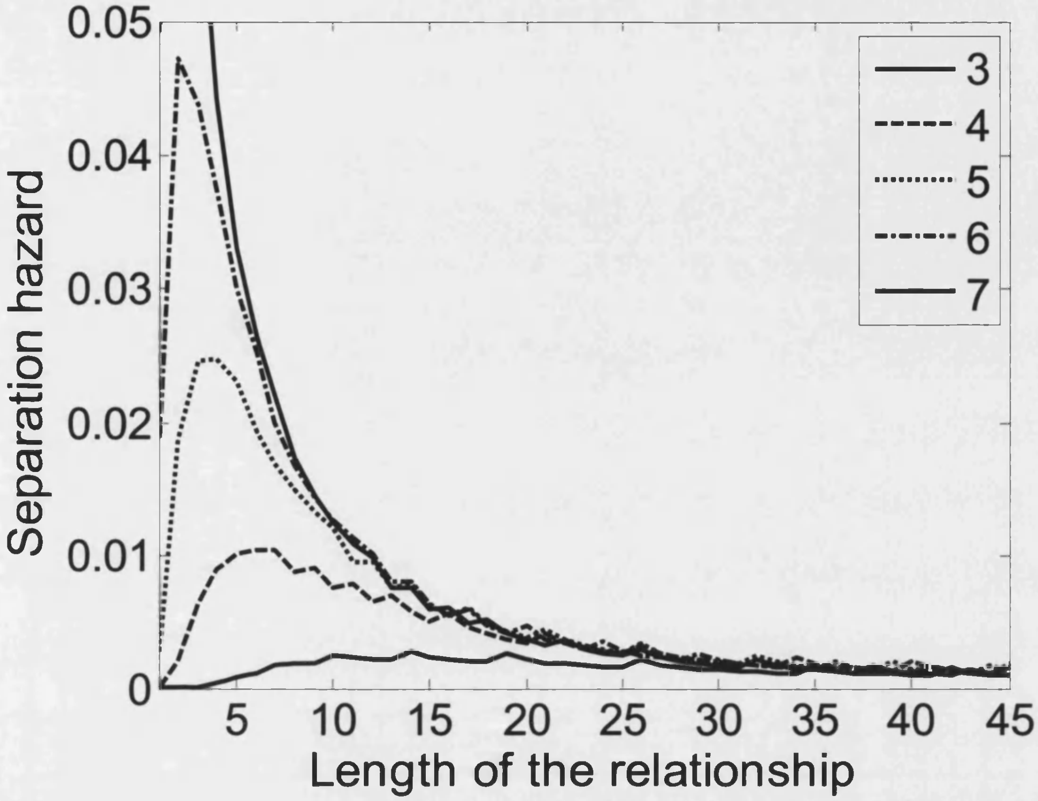

Notes: The figure shows the variance of the transition function for different prior standard deviations, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 21. Separation threshold for different prior standard deviations, normally distributed match qualities



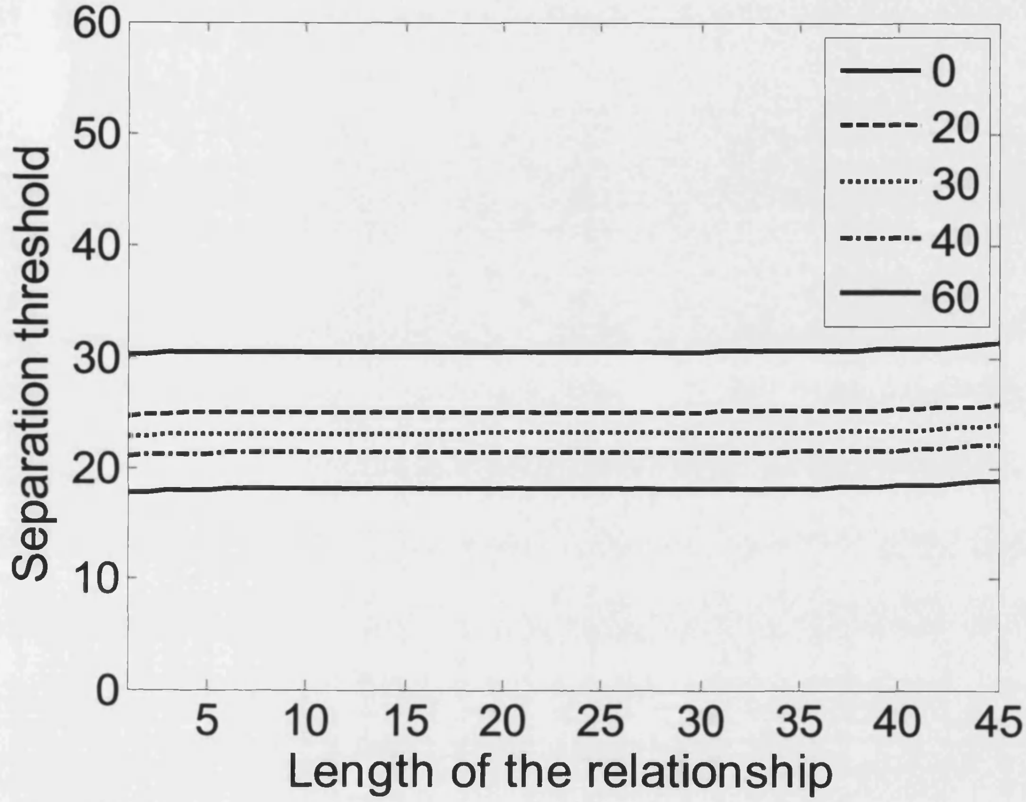

Notes: The figure shows the threshold for different prior standard deviations, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 22. Separation hazard for different prior standard deviations, normally distributed match qualities



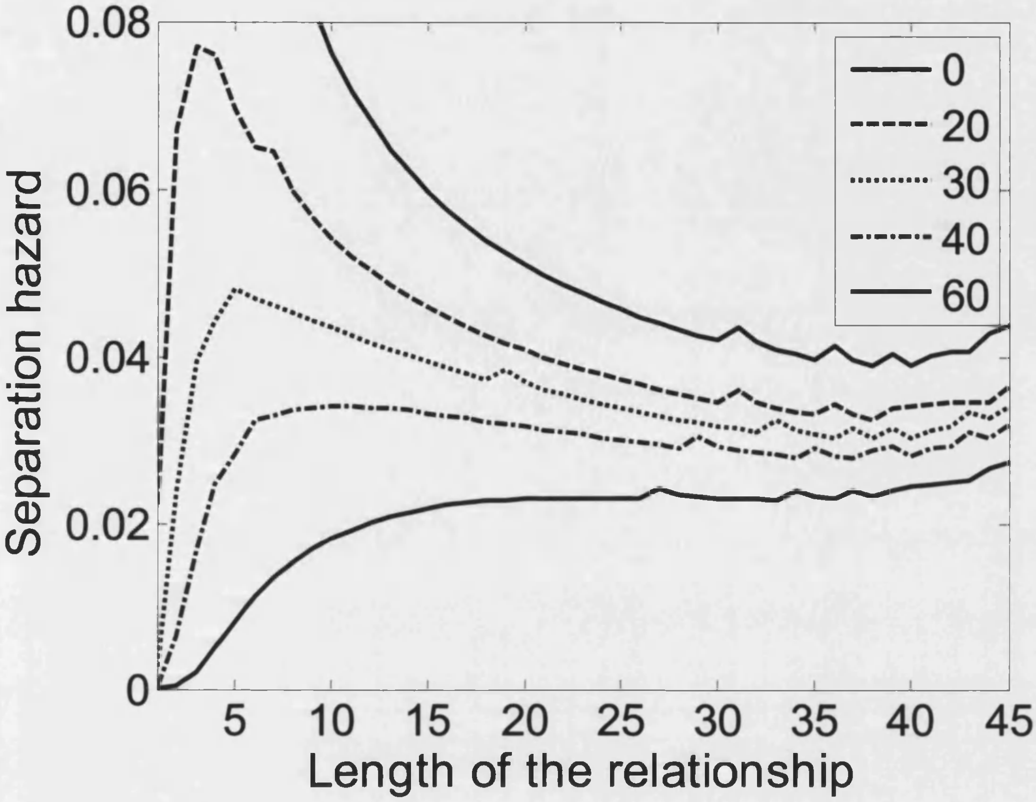

Notes: The figure shows the hazard for different prior standard deviations, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 23. Separation threshold for different separation costs, normally distributed match qualities and AR(1) process for match quality



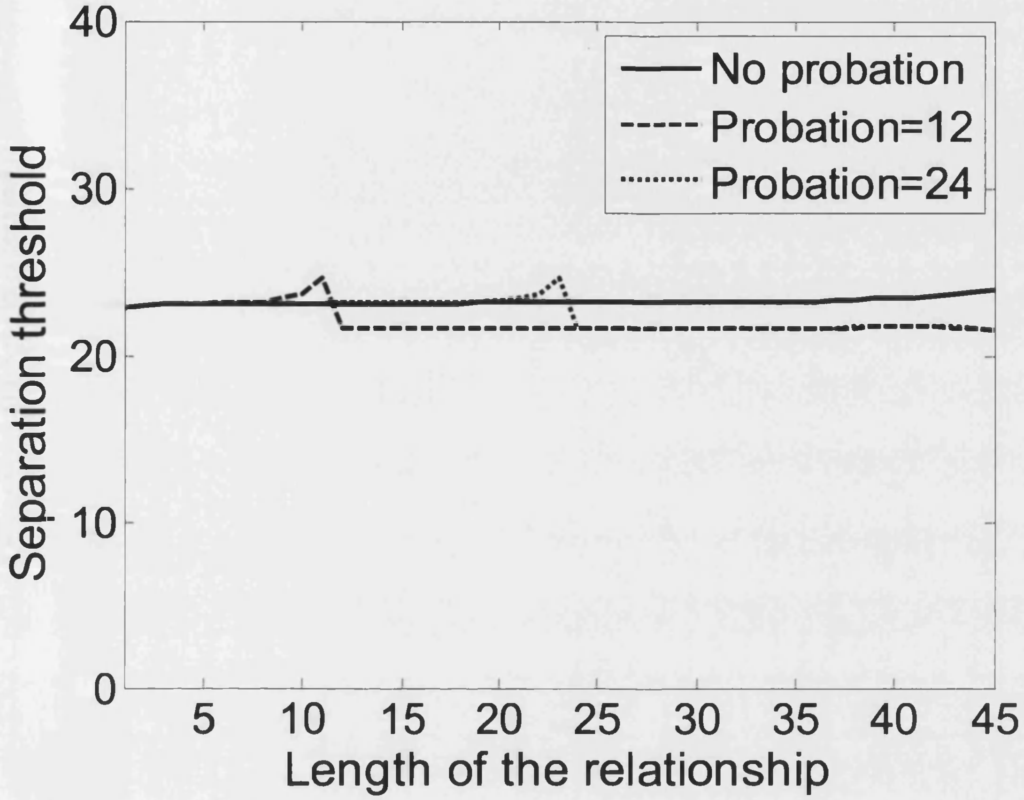

Notes: The figure shows the threshold for different separation costs, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 24. Separation hazard for different separation costs, normally distributed match qualities and AR(1) process for match quality



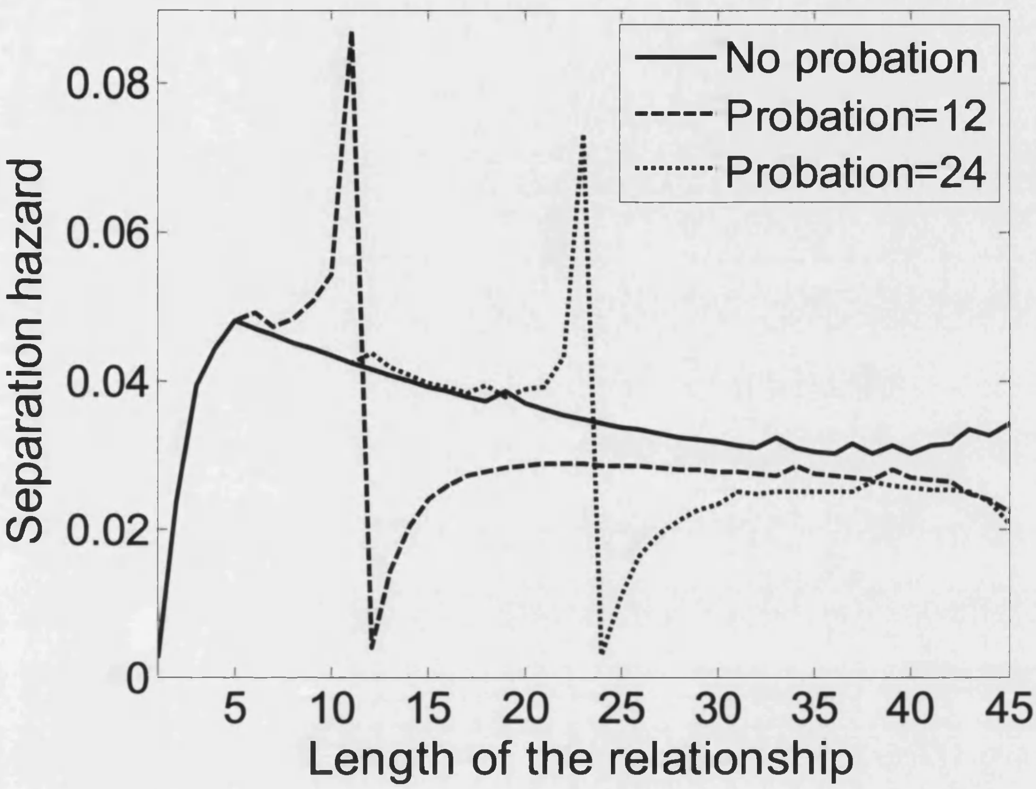

Notes: The figure shows the hazard for different separation costs, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 25. Separation threshold with a probation period, normally distributed match qualities and AR(1) process for match quality



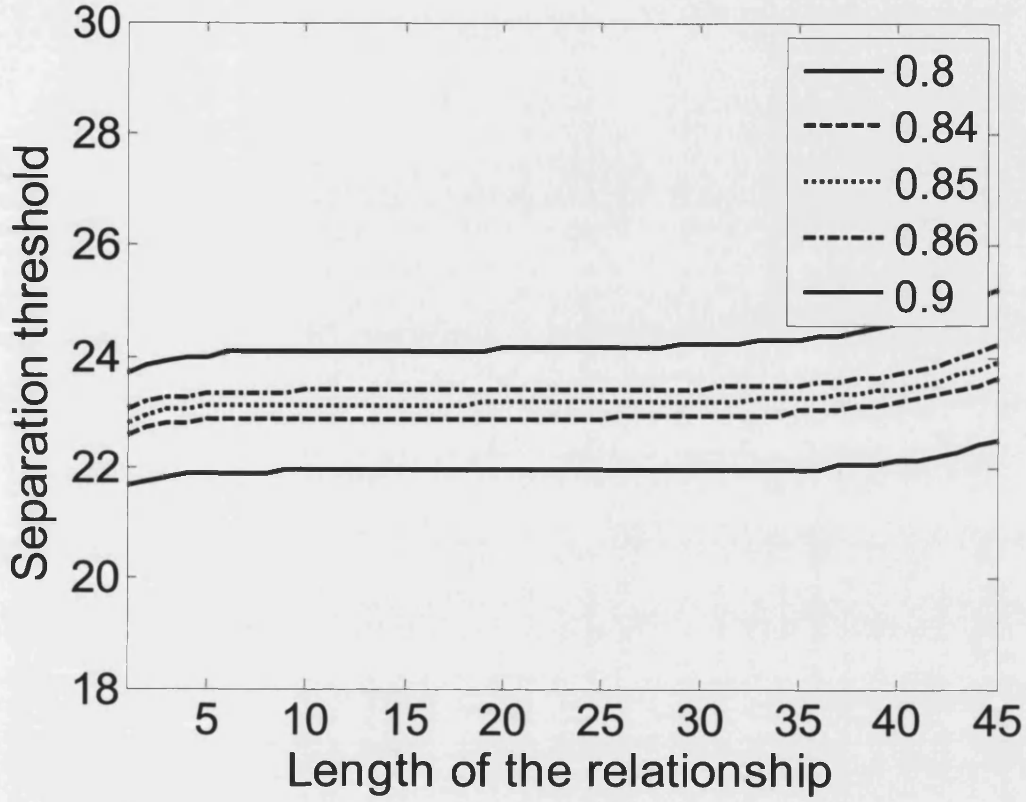

Notes: The figure shows the threshold for different lengths of the probation period, with separation costs being 1.5 during the probation period and 2.3 thereafter. All other parameters are fixed to their reference values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 26. Separation hazard with a probation period, normally distributed match qualities and AR(1) process for match quality



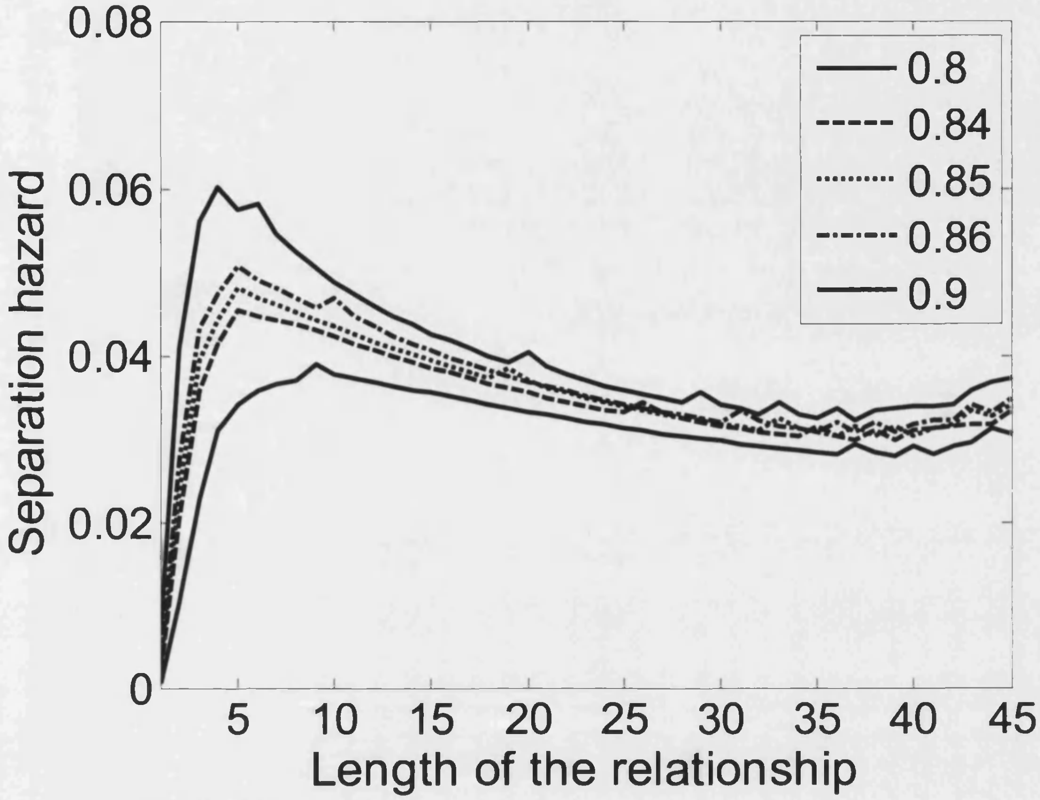

Notes: The figure shows the hazard for different lengths of the probation period, with separation costs being 1.5 during the probation period and 2.3 thereafter. All other parameters are fixed to their reference values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 27. Separation threshold for different discount factors, normally distributed match qualities and AR(1) process for match quality



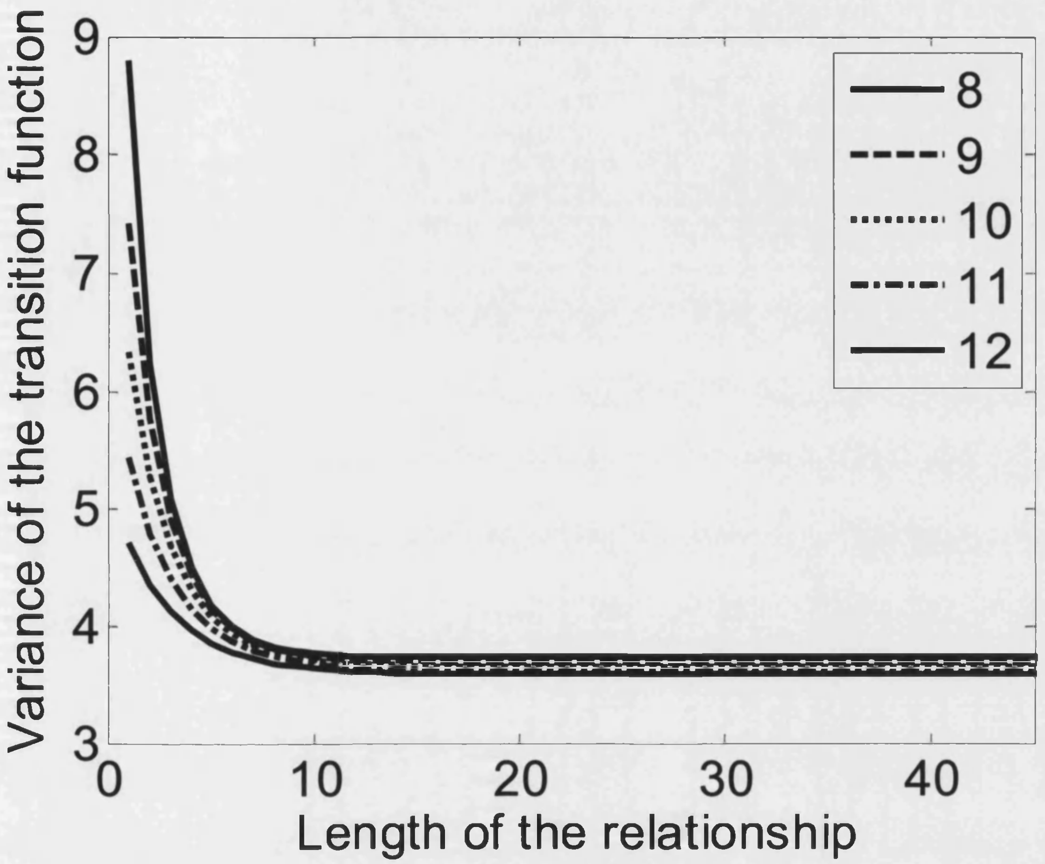

Notes: The figure shows the threshold for different discount factors, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 28. Separation hazard for different discount factors, normally distributed match qualities and AR(1) process for match quality



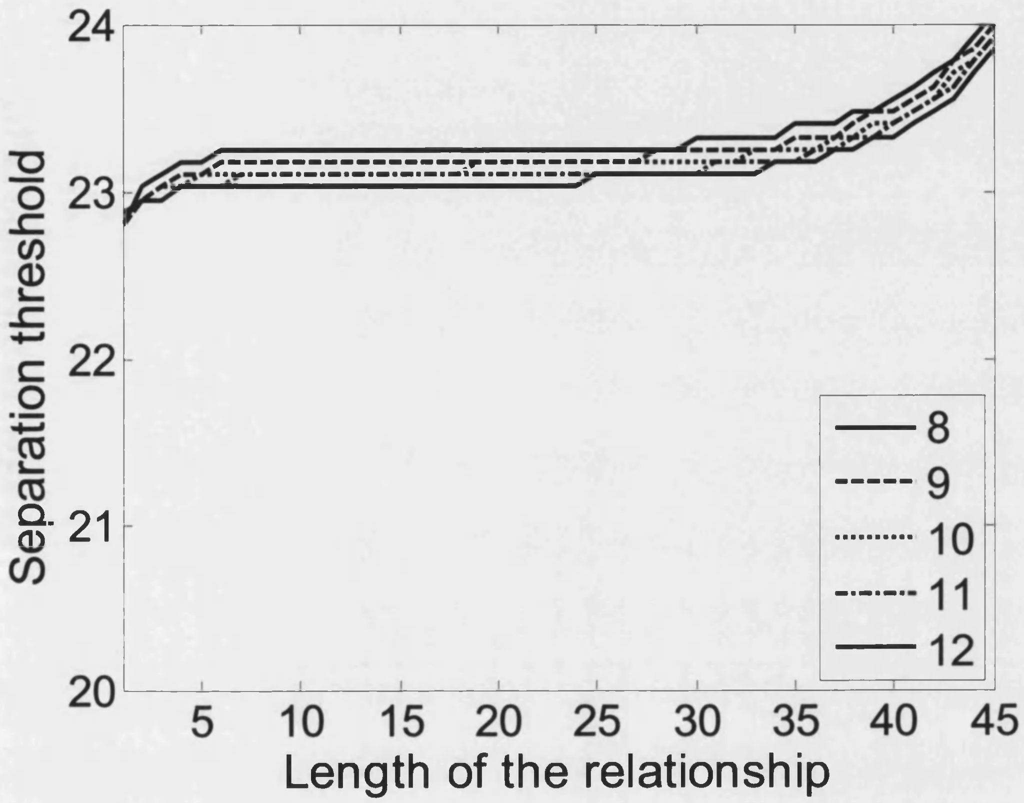

Notes: The figure shows the hazard for different discount factors, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 29. Variance of the transition function for different observation standard deviations, normally distributed match qualities and AR(1) process for match quality



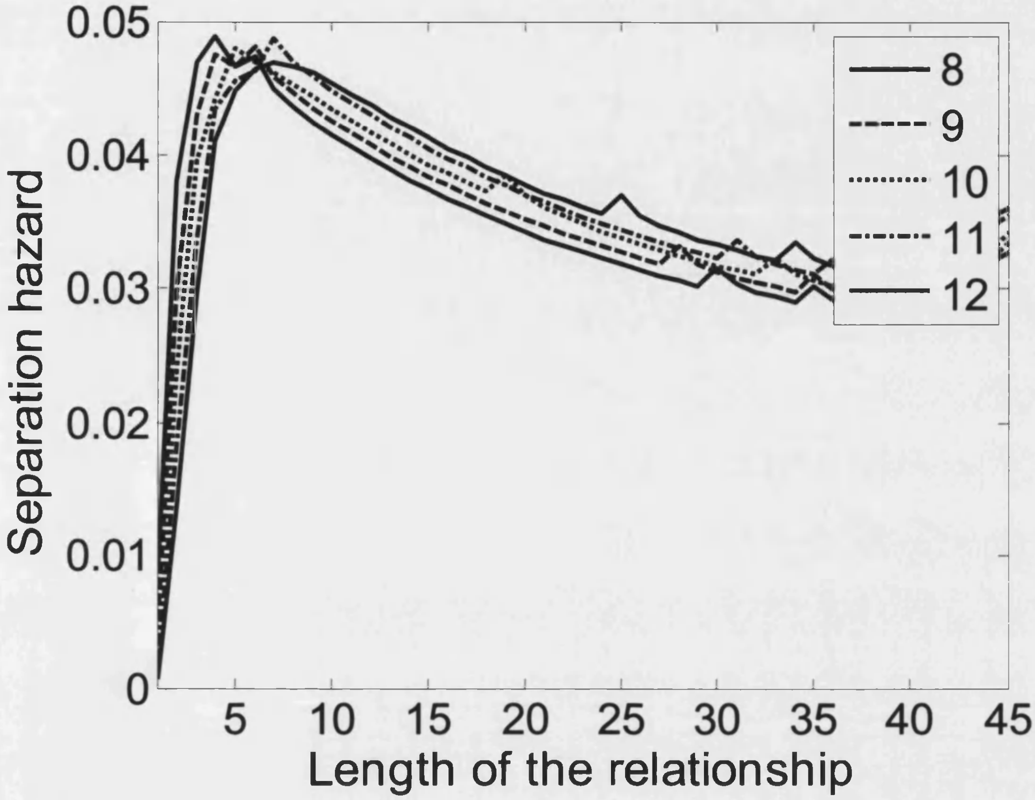

Notes: The figure shows the variance of the transition function for different observation standard deviations, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 30. Separation threshold for different observation standard deviations, normally distributed match qualities and AR(1) process for match quality



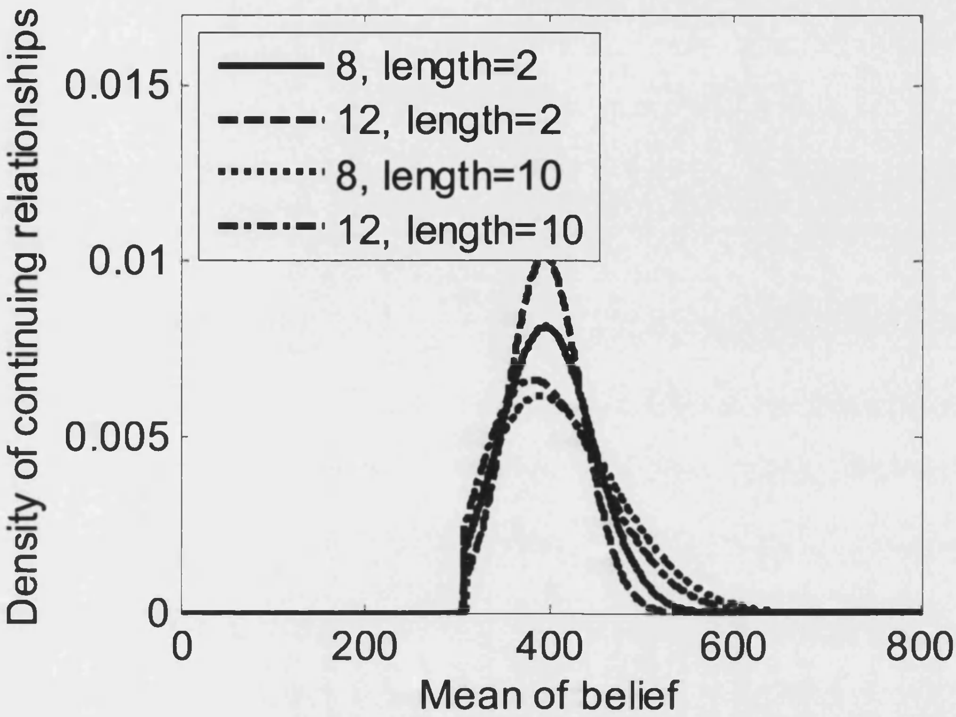

Notes: The figure shows the threshold for different observation standard deviations, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 31. Separation hazard for different observation standard deviations, normally distributed match qualities and AR(1) process for match quality



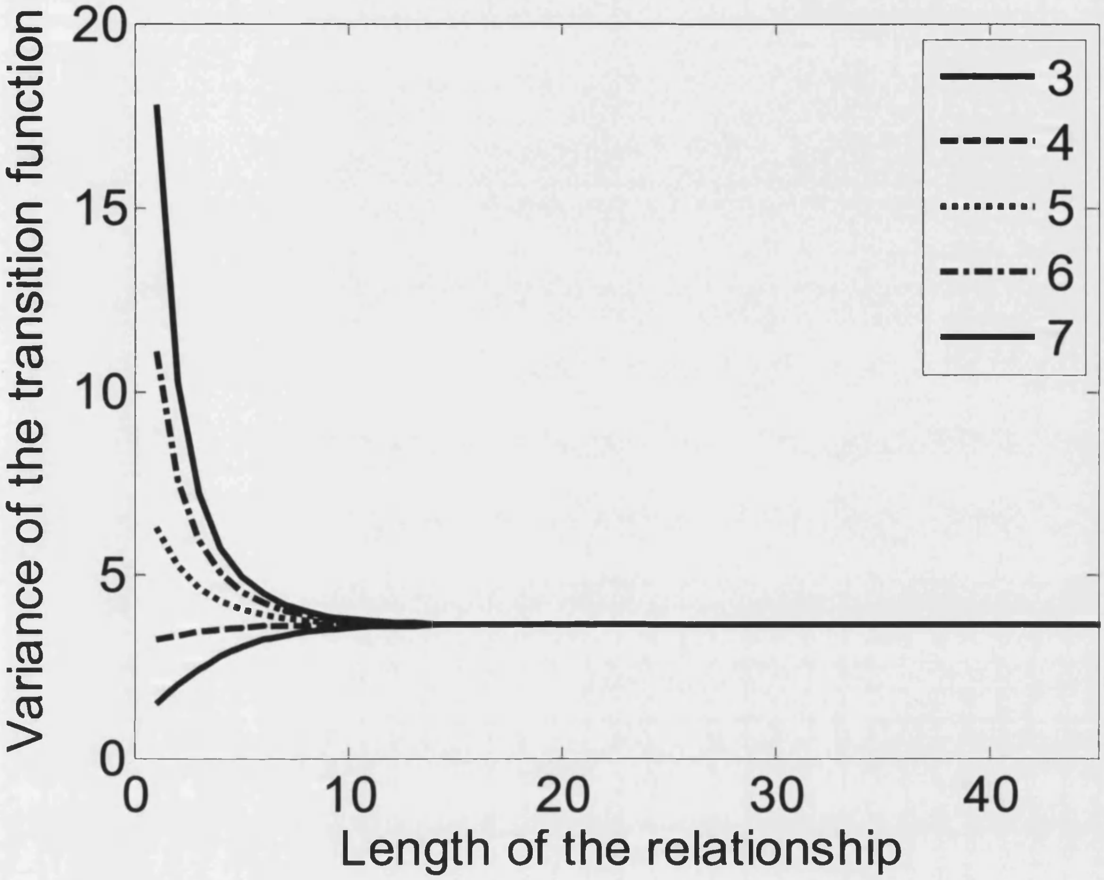

Notes: The figure shows the hazard for different observation standard deviations, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 32. Distribution of continuing relationships for different observation standard deviations, normally distributed match qualities and AR(1) process for match quality



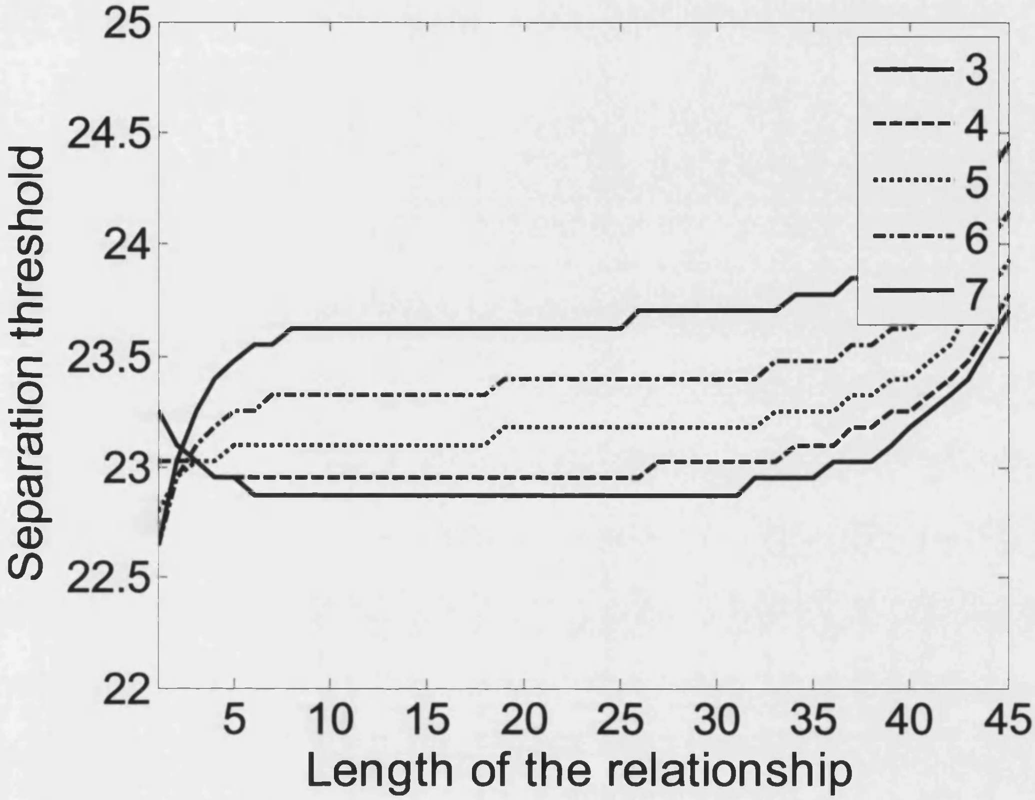

Notes: The figure shows the distribution of continuing relationships (i.e. the distribution $p_k(\hat{q}_k)$ in equation 2.28) for different observation standard deviations, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 33. Variance of the transition function for different prior standard deviations, normally distributed match qualities and AR(1) process for match quality



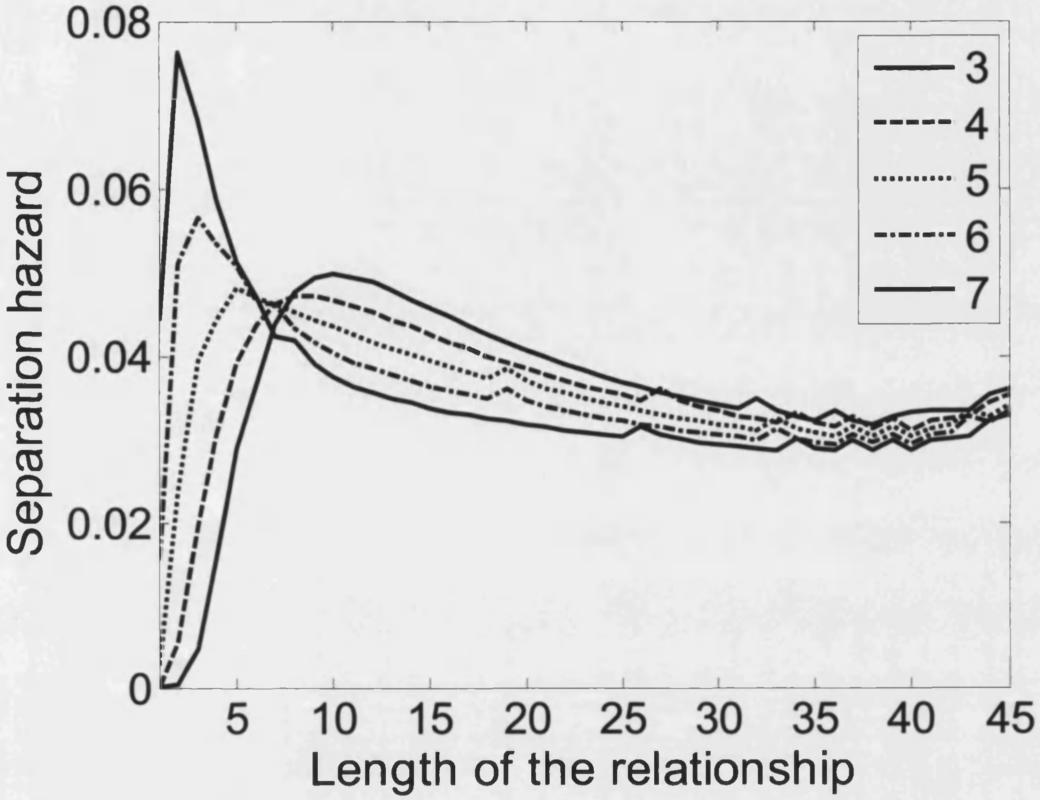

Notes: The figure shows the variance of the transition function for different prior standard deviations, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 34. Separation threshold for different prior standard deviations, normally distributed match qualities and AR(1) process for match quality



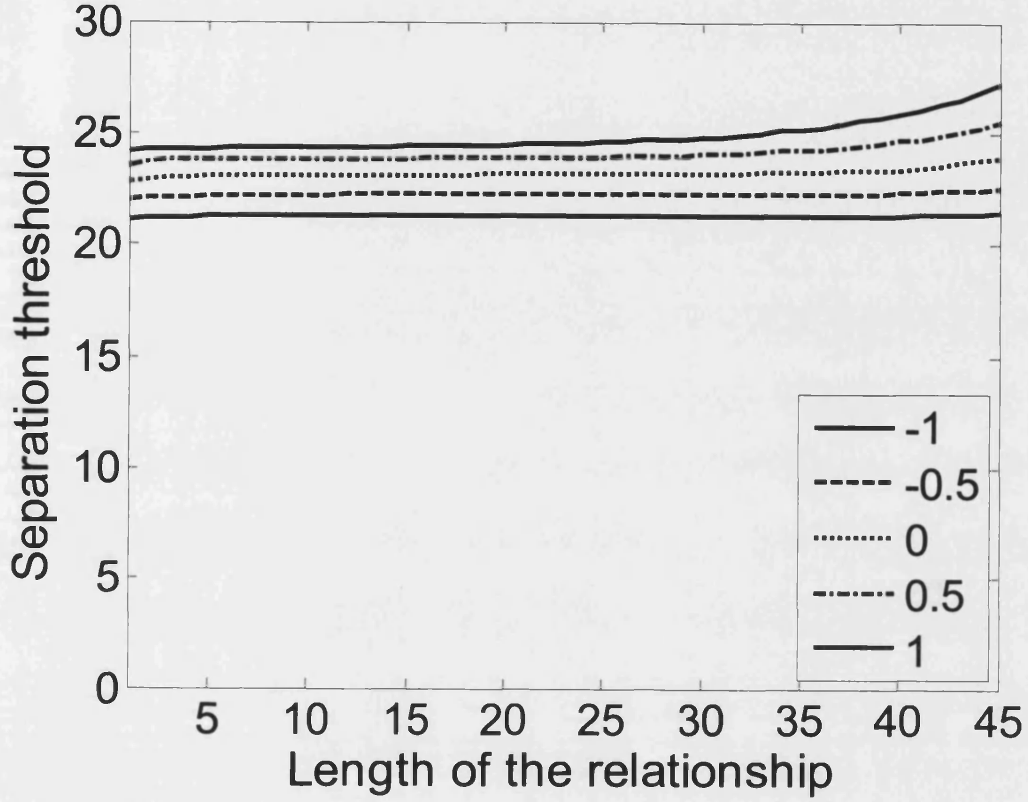

Notes: The figure shows the threshold for different prior standard deviations, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 35. Separation hazard for different prior standard deviations, normally distributed match qualities and AR(1) process for match quality



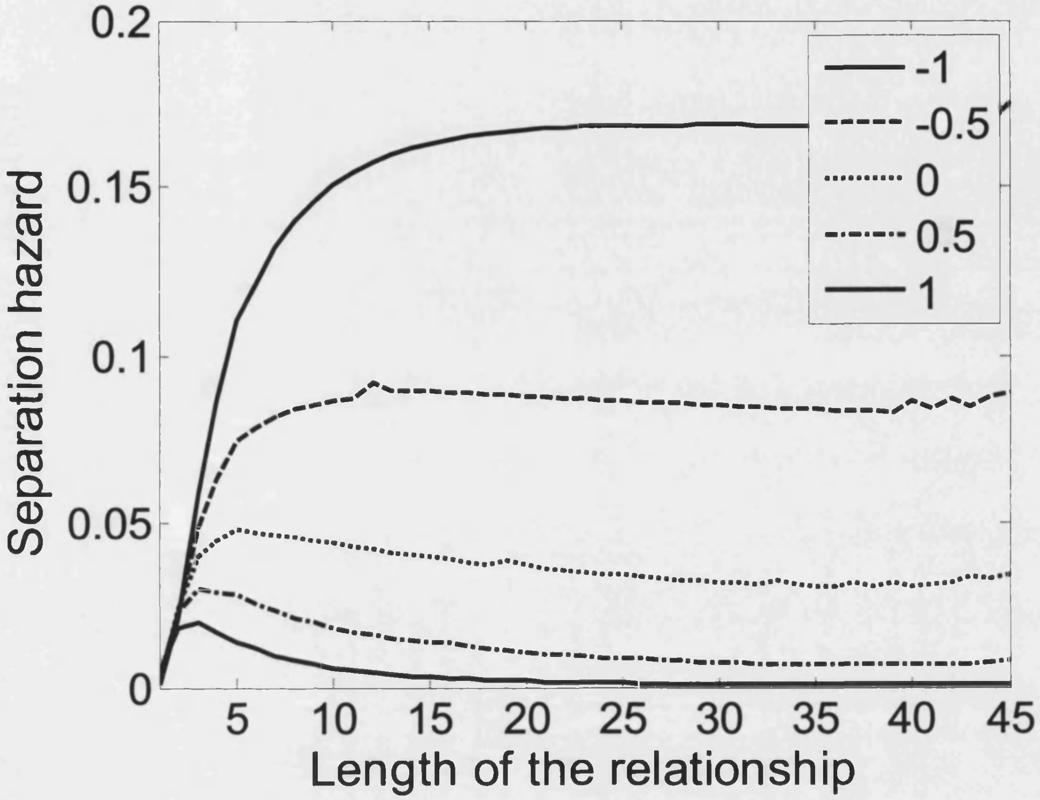

Notes: The figure shows the hazard for different prior standard deviations, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 36. Separation threshold for different process drifts, normally distributed match qualities and AR(1) process for match quality



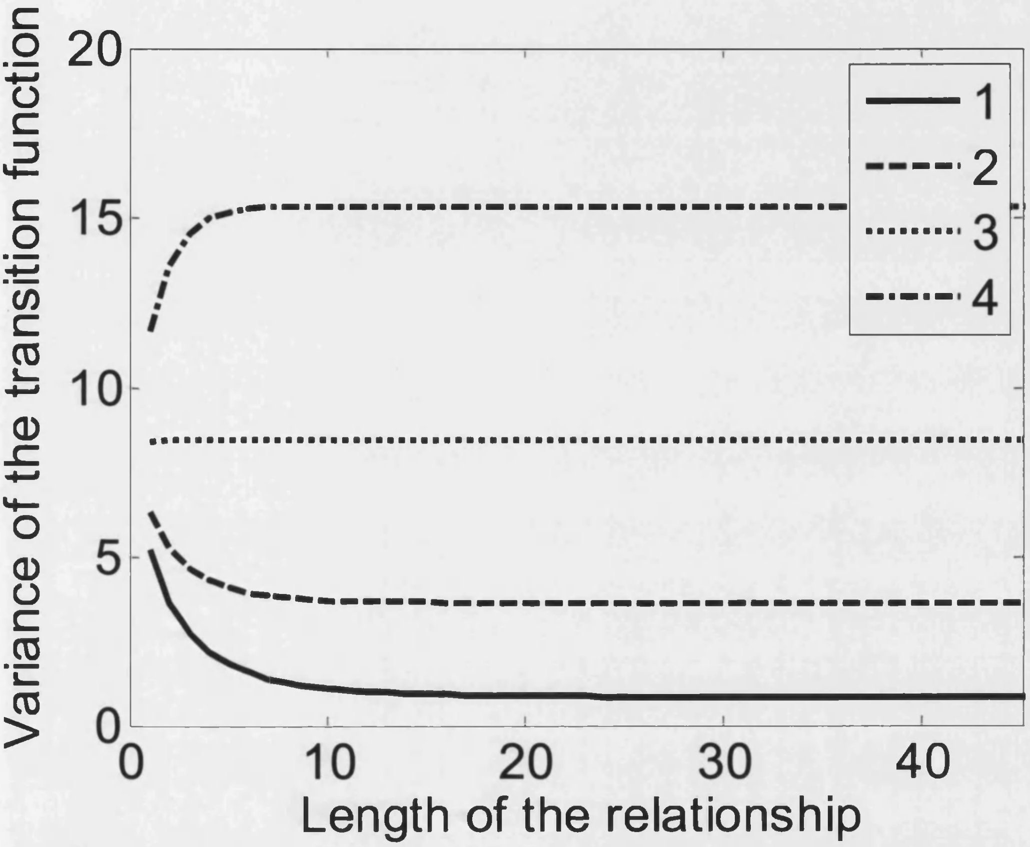

Notes: The figure shows the threshold for different process drifts, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 37. Separation hazard for different process drifts, normally distributed match qualities and AR(1) process for match quality



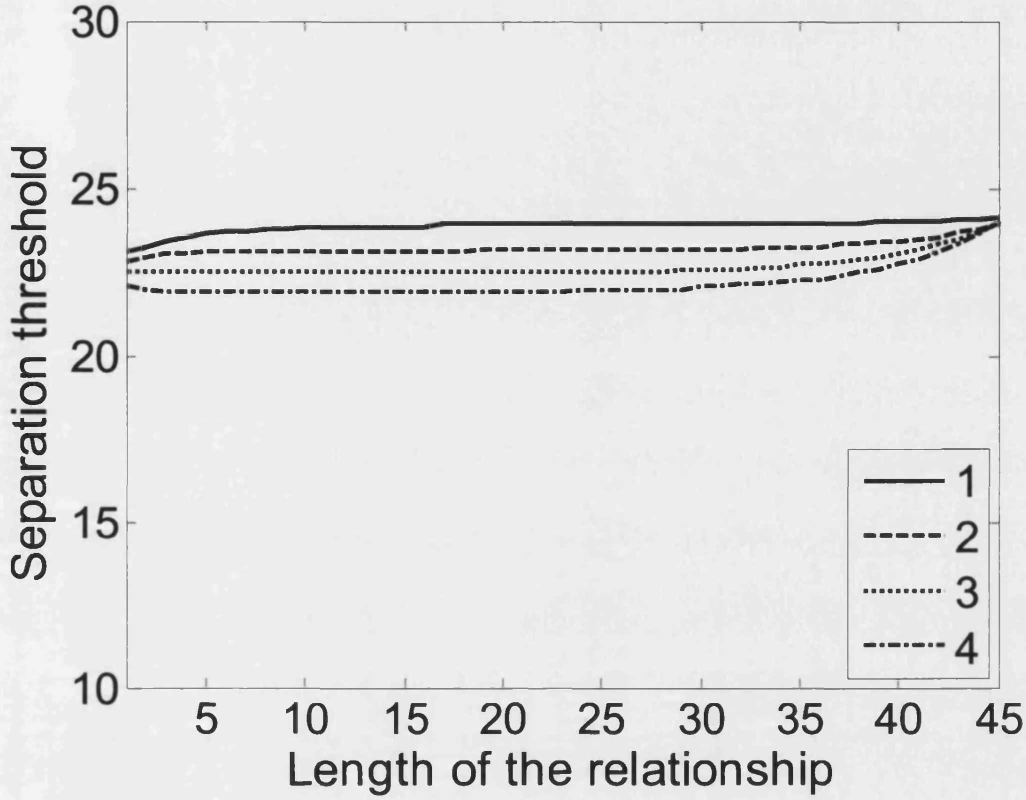

Notes: The figure shows the hazard for different process drifts, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 38. Variance of the transition function for different process standard deviations, normally distributed match qualities and AR(1) process for match quality



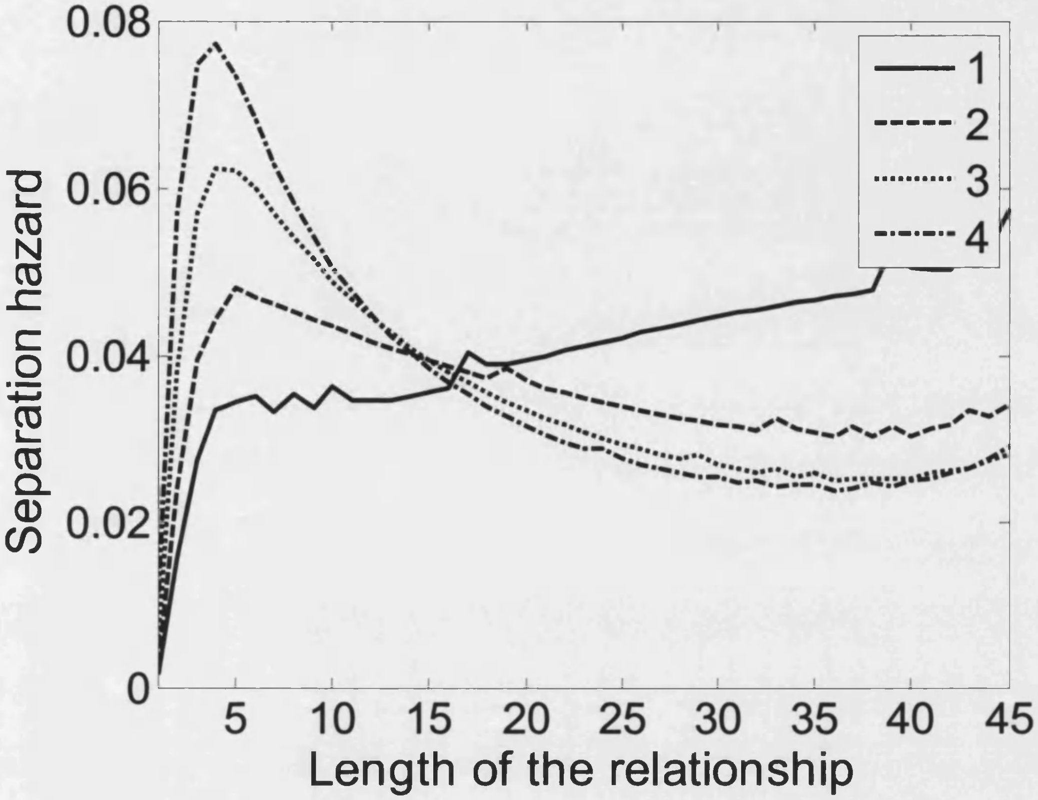

Notes: The figure shows the variance of the transition function for different process standard deviations, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 39. Separation threshold for different process standard deviations, normally distributed match qualities and AR(1) process for match quality



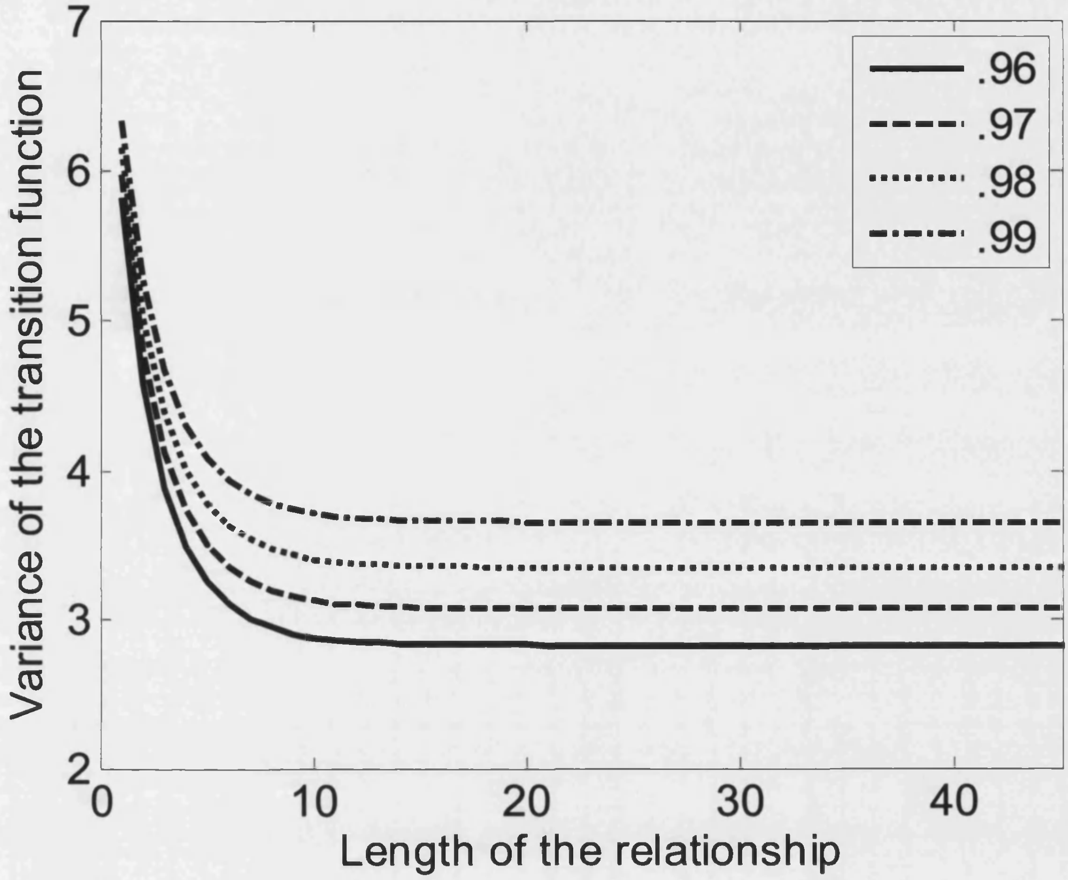

Notes: The figure shows the threshold for different process standard deviations, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 40. Separation hazard for different process standard deviations, normally distributed match qualities and AR(1) process for match quality



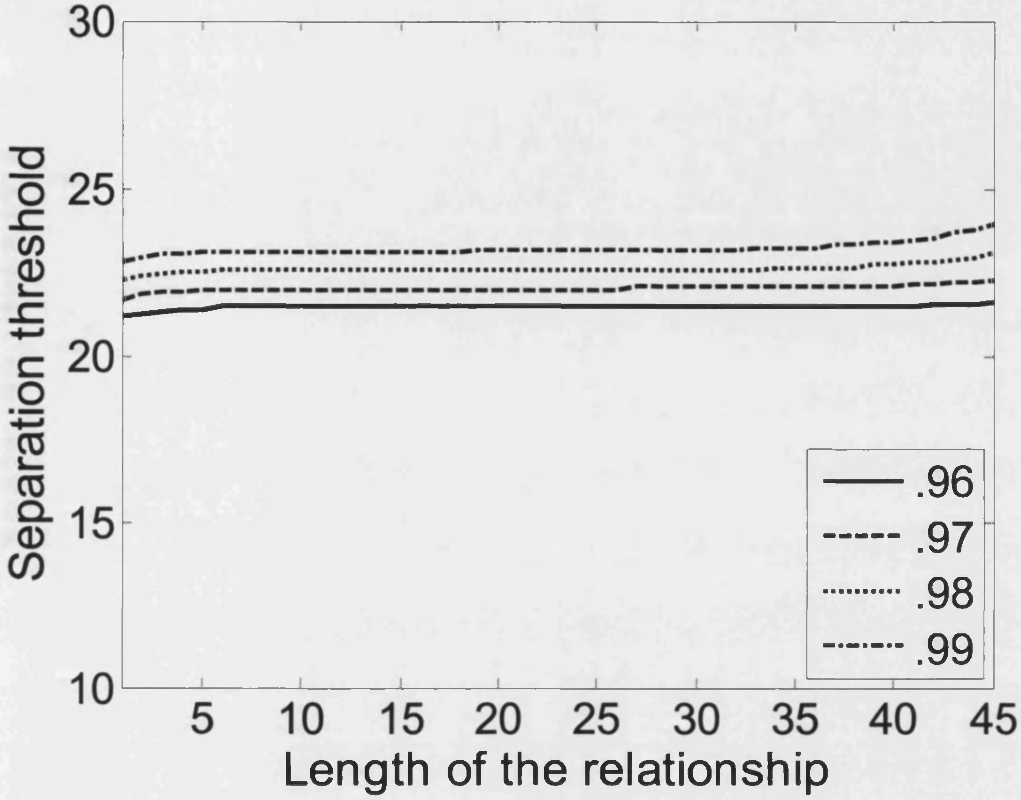

Notes: The figure shows the hazard for different process standard deviations, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 41. Variance of the transition function for different process AR(1) parameters, normally distributed match qualities and AR(1) process for match quality



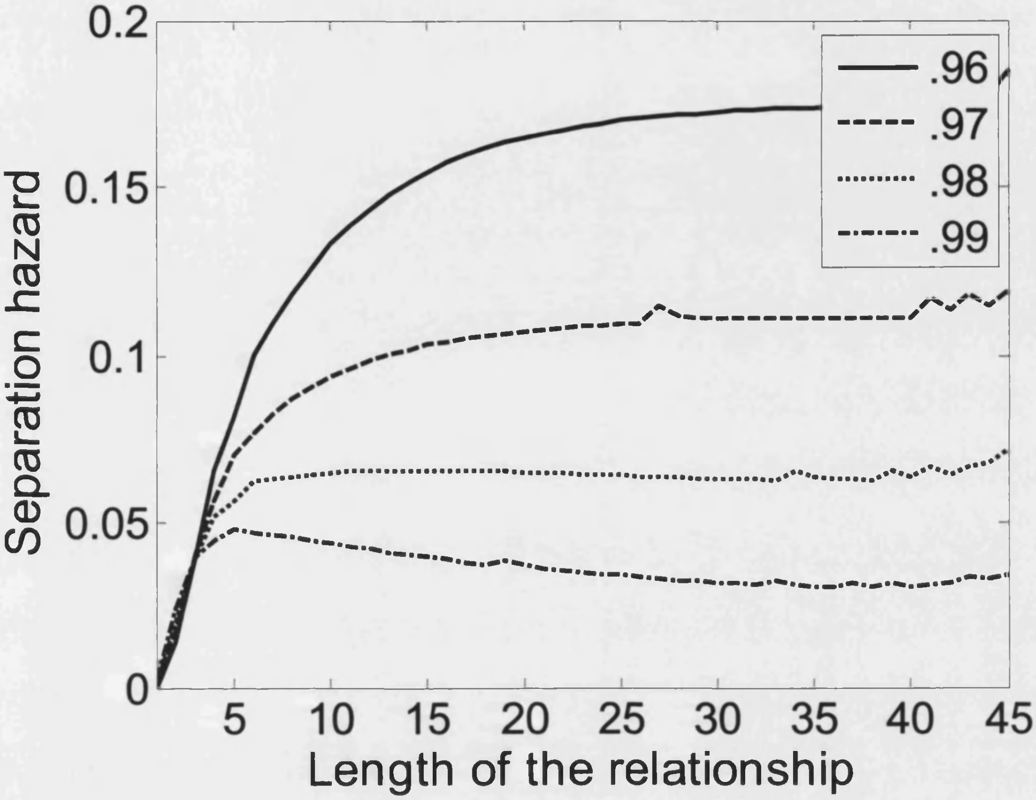

Notes: The figure shows the variance the transition function for different process AR(1) parameters, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 42. Separation threshold for different process AR(1) parameters, normally distributed match qualities and AR(1) process for match quality




Notes: The figure shows the threshold for different process AR(1) parameters, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

FIGURE 43. Separation hazard for different process AR(1) parameters, normally distributed match qualities and AR(1) process for match quality




Notes: The figure shows the hazard for different process AR(1) parameters, holding all other parameters fixed to their reference case values. Parameter values used in the reference case are to be found in Figure 1.

# Chapter 3

## Are Judges Sensitive to Economic Conditions? Evidence from UK Employment Tribunals[1]

*Ioana Marinescu*

### Abstract

In the view of classical legal theory, judges' decisions are fully determined by legal texts, whereas for legal realism and in particular law and economics, these decisions can be determined by other factors such as economic conditions. This paper specifically investigates whether judges deciding on the legitimacy of unfair dismissal claims are sensitive to economic conditions faced by firms and the workers they dismissed. Judges may face the following trade-off: in bad times, getting fired is more costly for workers, while at the same time firms find firing costs harder to bear. How do judges decide? I use British data on individual unfair dismissal and redundancy payment cases brought to Employment Tribunals in 1990-1992. Controlling for case selection, I find that when the unemployment or bankruptcy rate are high, *and* the dismissed worker has found a new job, judges tend to decide in favour of firms. All other things equal, when the dismissed worker is still unemployed, his probability of prevailing at trial is lower. However, the higher the unemployment rate, the more likely the unemployed dismissed worker is to win the case. On the whole population of cases brought to trial, a one point increase in the unemployment rate leads to a 7 points decrease in the probability of judges deciding in favour of dismissed employees. An increase in the bankruptcy rate has a similar effect. These findings are consistent with the idea that judges maximize the joint welfare of the dismissed worker and the firm, tailoring firing costs to local and individual economic circumstances.

[1] This paper was already included, in a very similar version, in my PhD defended in June 2005 at the EHESS, Paris.

"If a covenant be made, wherein neither of the parties perform presently, but trust one another; in the condition of mere nature (which is a condition of war of every man against every man), upon every reasonable suspicion, it is void: but if there be a common power set over them both, with right and force sufficient to compel performance, it is not void. For he that performeth first, has no assurance the other will perform after; because the bonds of words are too weak to bridle men's ambition, avarice, anger and other passions, without the fear of some coercive power [...]."

HOBBES, *Leviathan*, Part I, chapter XIV, §18.

# 1 Introduction

As Hobbes put it, contracts would be void without an enforcing power. From this basic requirement, legal theory usually goes a step further to posit that a contract is valid only if the parties freely agree to its terms. However, this basic requirement of contract law raises a double problem in the case of the contract of employment. First, the bargaining power of firms is usually higher than the bargaining power of individual workers, which casts doubt on the fairness of the contractual terms (A. Smith[2]). Second, the employment contract is generally incomplete, which gives rise to hold-up opportunities for both workers and firms (Malcomson, 1999). Although workers' shirking has often been stressed (Shapiro and Stiglitz, 1984), the subordination of workers to firms inherent in the employment contract opens large hold-up opportunities for firms as well (K. Marx[3]).

Labour law has developed to address these specific problems. One of the main areas of regulation concerns the conditions in which the employment contract can be terminated. Such regulation will be the focus of this paper. In the absence of specific regulation, employers and employees can terminate the employment contract at will, under some minimal conditions such as the requirement to act in good faith. Under regulation, the employer is typically required to have a good reason or just cause to terminate the contract. Thus, in most European countries, and sometimes in the United States[4], workers have the right not to be unfairly dismissed.

[2] *An Inquiry into the Nature and Causes of the Wealth of Nations*, Book I, chapter VIII, §11-13. Note that Smith believes that the imbalance in bargaining power should be corrected through the growth in the wealth of the nation, which increases the demand for labour, and not through a law fixing a fair wage (chapter VIII).

[3] *The Capital*, Book I, Section II, chapter VI, last paragraph.

[4] In the United States, this right is only granted in general if it is implicitly given by the employer. This is called the "implied-contract" exception to the doctrine of employment-at-will. For a study of the effects of this exception on employment, see Autor et al. (2002). In unionized firms, this right is

In as much as the legislator aims at protecting the workers against arbitrary job loss, unfair dismissal law can be viewed as having an ingrained "pro-worker" bent. However, this overlooks the autonomy of judicial bodies in charge of implementing the law. Judges have the possibility, more or less limited by each country's institutions, of tailoring the law to individual cases, according to their own views of fairness. Specifically, in countries such as the United Kingdom or France, judges in charge of implementing unfair dismissal legislation are themselves employees and employers meant to represent their respective constituencies[5], which makes them particularly sensitive to the specific context of the case. For example, economic circumstances affect the costs incurred by firms and workers when a dismissal takes place. It is more difficult for dismissed workers to find a job in a high unemployment context. For a firm, being over-manned is more hazardous when bankruptcy risk is higher. Testing whether and how economic circumstances may tilt the sense of legitimacy of judges in the marginal unfair dismissal case is the main goal of this paper.

However, the exact effect of economic conditions on judges' decisions depends on the definition of their objective function. Assuming for example that judges' objective is to maximize social welfare, it is not clear how their decisions should be related to economic conditions. Indeed, the right to claim unfair dismissal gives rise to a firing cost incurred by firms whenever a dismissed employee goes to court. If such a cost discourages firing, it also discourages hiring so that the effect on employment is unclear. Non surprisingly then, economic models show no clear-cut relationship between the level of firing costs and the level of employment (Bertola(1992) in partial equilibrium, Ljungqvist(2002) in general equilibrium). The empirical literature using cross-country variation does not reach a clearer conclusion. Djankov et al. (2003) find that a general employment law index has a positive correlation with the unemployment rate, i.e. a more protective employment law is correlated with higher unemployment. Within the OECD however, Employment Protection Legislation (EPL) is found to have no significant relationship with the unemployment rate (OECD, 1999). It is then interesting to study what judges' revealed preferences tell us on their views on fairness as a function of economic conditions. Judges may try to

granted by the contract between the union and the firm, which specifies some rules to guarantee the fairness of the discharge or lay-off process.

[5] In the UK, the body of judges also includes a professional judge.

restore efficiency where the blind enforcement of the law is inefficient and private contracting is not an option. Indeed, some dispositions of dismissal law severely limit the ability of agents to defeat it using private arrangements, which largely invalidates the view advanced by Lazear(1990) according to which the requirement of a severance payment could be undone by properly designed contracts.

The literature on the effect of macroeconomic conditions on EPL enforcement by labour courts is scarce. Using regional aggregated data, Macis (2001) finds a negative effect of the unemployment rate on the share of employees winning their unfair dismissal cases. Ichino et al. (2003) use micro data from a large Italian bank (86 trials in 15 years) combined with Macis' macro data and find a positive effect of the unemployment rate on the probability of an employee winning the unfair dismissal case.

This study uses a survey of British Employment Tribunal cases to estimate the effect of economic conditions on workers' probability of winning unfair dismissal cases at trial, and hence determine whether judges' decisions are influenced by economic conditions. The determination of the relevant empirical strategy faces however an important caveat. Indeed, an abundant literature (Cooter and Rubinfeld,1989) has pointed out that the distribution of case quality at trial can be expected to be different from the distribution of case quality in the population of applicants. In my paper, *case quality* is conventionally defined as being the quality of the *worker's* case, i.e. a higher case quality means that the worker would, all other things equal, have a higher probability of winning at trial. The focus of this paper is then to estimate the impact of economic conditions on judges' decisions. However, given the above mentioned problem, estimating this effect using only the sample of cases having reached the trial stage may be misleading. Indeed, it could be that economic conditions are correlated with case quality. In particular, one concern may be that unemployed workers are more likely to go to trial, thereby reducing the average quality of cases being put forward by workers. This would generate a negative correlation between unemployment and the quality of the cases going to trial. Note that the employment status of the dismissed worker determines the existence of a bias: if the worker finds a new job right after being dismissed, the unemployment rate is unlikely to influence his decision to go to trial. The data used in this paper contains information on the employment status of the worker and thus allows one to estimate the effect of economic conditions on judges' decisions on a category of cases for which contamination by selection bias is unlikely.

In general, if case quality can be measured sufficiently well, any effect of economic conditions on case quality is captured by the case quality measure, and therefore the estimated effect of economic conditions corresponds to the effect on judges' decisions. Like Ichino et al, I have information on the reason for the dismissal, as well as other individual variables that may be correlated with case quality, such as, crucially, the amount of settlement offers made by firms to workers prior to trial. Thus, the measures of case quality in the data are exceptionally good by the standards of the literature. If these measures were still not precise enough, one would need to account for potential selection on unobservables. I carefully analyse the selection of the sample of applicants itself, as well as the selection of applicants for trial; the latter analysis is performed using sample selection models by Heckman(1979), Van de Ven and Van Praag (1981), and Sartori(2003). In all the models considered, I find a negative effect of both the unemployment rate and the bankruptcy rate on workers' probability of winning their cases, rejecting the possibility that the main results of this paper are driven by selection bias. When also controlling for the worker's employment status, I find that workers who have found a new job see a decrease in their probability of winning when unemployment or bankruptcy rates are higher, whereas workers who are still unemployed see a *positive* effect of the unemployment rate on their probability of winning.

Economic conditions thus affect judges' decisions differently for employed and unemployed workers, while on average, worse economic conditions make judges marginally more pro-firm. This result should be taken into account by legislators when framing unfair dismissal legislation: indeed, the effect of the law is a combination of the formal content of the law and the way judges actually enforce it. For social scientists, the finding of this paper indicates that one should take into account enforcement when assessing the efficiency of unfair dismissal legislation, and EPL in general. Indeed, what really matters for economic performance, and therefore economic policy, is not EPL per se but the effective firing costs induced by its application.

The paper is structured as follows. Section 2 gives some background on British Employment Tribunals and describes the data used. Section 3 discusses theories of judges' decisions and defines the estimation problem arising due to sample selection.

Section 4 deals with the selection of the sample of applicants, establishing that there are no observable effects of economic conditions on applicants' case quality. Section 5 presents a general model of settlement behaviour and derives the relevant econometric specifications. Section 6 gives the results of the empirical analysis. And section 7 concludes.

# 2 British Employment Tribunals and data used

## 2.1 British Employment Tribunals and the employment law

Most European countries have specialized labour tribunals to deal with unfair dismissal cases, and other specific labour law cases that may arise. It is commonly assumed that dealing with these matters requires some knowledge of common practices among firms and workers. Some countries, such as France and the United Kingdom, have decided it is in the best interest of equity to have representatives of employees and employers act as judges and provide the expertise required. In the United Kingdom, the employment tribunal is composed of one chairperson, a professional judge, and two appointed lay judges, one representing employers and the other representing employees. The lay judges are chosen by the administration from lists of persons proposed mainly by trade unions (for lay judges representing employees) and employer groups (for lay judges representing employers).

The United States have no such specific labour courts, but the Employment Tribunals' setting in the United Kingdom is similar to the arbitration scheme used in unionized firms in the United States to decide on issues where employer and union disagree (Ashenfelter and Bloom 1984, Farber and Bazerman 1986). In both cases, the institutional setting is meant to achieve some equitable compromise between firms' and workers' interests. In an experimental study, Farber and Bazerman (1986) find that when deciding on a wage increase, the arbitrator reacts in an asymmetric way to firms' financial situation. Compared to a medium situation, worse financial conditions lead to a discount in the award made by the arbitrator and better financial conditions lead to a premium. Interestingly enough, the premium is significantly lower than the discount. This shows that arbitrators are particularly sensitive to firms' interests in bad times, and suggests that judges in labour courts may react in a similar fashion.

In Europe, the majority of cases labour courts have to deal with concerns dismissals. In the US, although the economics literature has focused on arbitration on wages issues, these are only a very small part of the issues arbitrators have to decide on. Instead, issues of discharge and disciplinary action are most common (see for example the statistics given by the Federal Mediation and Conciliation Service, www.fmcs.gov). In other terms, in the arbitration system, dismissal is at the centre of debate, just as in British Employment Tribunals. In what follows, I am going to concentrate on cases concerning dismissal, although I also have data on other types of cases such as unfair deduction from wages, and race and sex discrimination[6].

Once he/she has been dismissed, the employee can bring a case to court, either to ask for some severance/redundancy payments if those are absent or insufficient, or to ask for compensation for unfair dismissal. It is important to notice that the first category of cases (redundancy and severance payments) is closer to the second one (unfair dismissal) than it may seem at first glance. Indeed, if the employer claims very serious misconduct on the part of the employee, then the employer need not pay any severance payment to the employee. In those cases, the employee, without claiming there was no reasonable ground for his/her dismissal, can still claim that the misconduct was not as severe as to deprive him/her of a severance payment; this is then very close to saying that the dismissal was in some way unfair. A surprising but fruitful parallel can be drawn here with unemployment insurance. Indeed, in many countries, and in particular in the United States, workers do not receive any unemployment benefits if they have lost their job by their own fault. This restriction has particularly interesting consequences in the United States where experience rating is in place. If a firm wants to avoid paying higher unemployment insurance contributions when laying off more often (this is the principle of the so-called "experience rating"), it can instead discharge its employees for misconduct or underperformance, which has no effect on its experience rating. Of course, such opportunistic behaviour should be limited: this is why fired workers can appeal against their disqualification for unemployment benefits by showing that they did not misbehave or shirk in such a way that the firm could have legitimately discharged them. If the unemployment insurance commission decides in the worker's favour, the worker receives unemployment benefits and the firm does get penalized in its

[6] I tested for an the influence of economic conditions on those other cases and found similar but less significant effects. The sample however is too small to provide reliable results.

experience rating. Thus, appealing against disqualification to the unemployment insurance commission in the United States is similar to filing a case for unfair dismissal in Europe.

Returning to the British Employment Tribunals, the British law governing unfair dismissal cases is formulated in such a way that it explicitly allows judges to take into account circumstances other than the mere facts pertaining to the case (the "substantial merits of the case"):

> "the determination of whether the dismissal was fair or unfair, having regard to the reason shown by the employer, shall depend on whether in the circumstances (including the size and the administrative resources of the employer's undertaking), the employer acted reasonably or unreasonably in treating it as sufficient reason for dismissing the employee; and that question shall be determined in accordance with equity and the substantial merits of the case." (Employment Protection (Consolidation) Act 1978 s. 57(3), as amended by Employment Act 1980, s. 6).

Thus, three elements can legally determine judges' decisions. First, as is obvious, decisions shall depend on the substantial merits of the case. But decisions may also depend on relevant "circumstances", which explicitly include the firms' ability to deal with such cases in a proper, and hence costly, way. The implication is that smaller firms, and firms with less administrative resources, may expect more forgiveness on the part of the judges. Moreover, the list of circumstances is not explicitly limited and therefore economic conditions could also in principle be included, as firing costs are more difficult to bear for firms when economic conditions are worse. Decisions shall also depend on "equity", which means that judges should compromise between firms' and workers' interests.

To see how these considerations apply to a specific case, we can take an example from a 2003 Employment Tribunal decision concerning the allegedly unfair dismissal of a truck driver. During an early delivery up a particularly tricky lane, the truck ended up on its side, resulting in damage to vehicle and interruption of deliveries that day. The employer observed the scene and, without further inquiry, dismissed the driver without notice for gross misconduct ("reckless driving"). The employer argued that this was gross misconduct as it was a financial disaster for his business: he could not afford to increase insurance premiums by claiming on the insurance policy for the damage to this vehicle. The driver, who had by then found a new job, argued this was only an unfortunate accident, and that such things happened in the past without the driver being dismissed. The court decided for the driver, mainly on account of the fact that the employer had not followed the rules set out in the company's own handbook, according to which no dismissal should take place without a reasonable investigation and an opportunity for the employee to offer an explanation.

This example calls for two comments. First, deciding whether the employee was guilty of a gross misconduct partially depended on judges' view about the fact that the employee's misconduct was endangering the financial position of the firm; hence, if economic conditions were bad, the argument of the employer would sound more credible. Second, the decision depends on procedural fairness: the employer not respecting a certain rule of conduct was seen as a fault. This is very similar to the American implied contract exception, whereby it is insisted that if the employer stated, even implicitly, that the employee has a right not to be unfairly dismissed, then the employer may be found guilty of unfair dismissal at court. In the British context, one must remember that one of the reasons why the unfair dismissal legislation was introduced in the first place was to give an incentive to employers to organize a systematic internal disciplinary procedure to deal with conflicts arising at the workplace (Davies and Freedland, 1993).

At this point, one might wonder whether judges are allowed to give their own interpretation of the fairness of the dismissal based on considerations such as economic conditions. More precisely, even though the law in its formulation may allow such considerations to have an influence on decisions, it may be that the appeal courts do not allow it. However, the Court of appeal decision in the Gilham and others v. Kent County Council case in 1985 leaves tribunals full discretion to decide on matters of facts. In the British legal system, the control of the court of appeal only concerns breaches in the principle of the law itself, and no appeal on matters of fact can be made. The way the Gilham case arose deserves some further comments. In the early eighties, the conservative government of Mrs. Thatcher cut local authorities' budgets. The Kent County decided to reduce dinner ladies' (the persons working at school restaurants) wages to face the new financial constraint. Some dinner ladies refused this modification to their conditions of employment, were fired and claimed unfair dismissal. The court balanced employers' financial constraints and the fact that there had been a breach in a nationally negotiated agreement concerning wages, and decided in favour of the dinner ladies. An appeal was made by the employer. The employer's lawyer argued that a pro-employer decision was reached in two quasi-identical cases in Devon and Somerset counties. The court of appeal confirmed the tribunal's decision, stating that different courts are permitted to come to different conclusions in similar cases: "Now whether or not an employer has behaved reasonably in dismissing an employee is a question of fact, and it is a question upon which different people, looking at the same set of circumstances, may reasonably come to different conclusions. It is therefore endemic in a system where there is no appeal on fact [because of the high costs it would involve] that from time to time different industrial tribunals will give different answers to broadly similar situations [...]". In this specific case, one should note that the two pro-employers decisions cited by the Kent County's lawyer were taken in a high-unemployment (7.2%) region, whereas the pro-employees decision in the Kent county was taken in a lower unemployment (6%) region. Of course, as these decisions concern public sector firms, one cannot argue that judges give more weight to firms' arguments in bad times *because* these particular firms may go bankrupt. However, it can still be that judges are more sensitive to pro-firms arguments in general when economic conditions are worse. This can be either because of mere association of ideas or because the sense of fairness of treatment for workers is related to economics conditions. In the first case, by association of ideas, firms in the public sector end up being given a similar treatment to firms in the private sector although they do not to have the same financial constraints. In the second case, judges may consider it less legitimate for the dinner ladies to complain about a change in their wages and getting dismissed on that account when the local economy is undergoing some episode of relatively bad economic conditions. One can picture the judges using reasoning such as: "these dinner ladies should be content to have a job, and they should be happy to make some sacrifices to keep it, as so many other workers are unemployed and so many firms are under heightened financial pressure".

To conclude, the functioning of the British dismissal law allows Employment Tribunals judges to take into account economic conditions when deciding on whether or not a firm has acted reasonably in dismissing an employee. I will now describe the data used to investigate whether this is indeed the case.

## 2.2 Data used

I have data on individual cases, coming from the 1992 survey of Employment Tribunal Applications in Great Britain. This survey was conducted in the following way. First, a random sample of applications completed between January 1990 and October 1991 was drawn; then, employers and employees involved in those cases were interviewed. Note however, that, to save on resources, the survey managers decided to interview all employers and only half of the dismissed employees involved in the cases of the sample. The sample is constructed to be representative of all cases, withdrawn, settled or heard. Many variables are available, including the precise reason for dismissal, and information on all the stages of the case from application to tribunal hearing, including details of settlements, such as the amounts firms offered to workers for a settlement.

Among the available variables, I pick a set $X$ that will constitute the control variables: they are variables concerning case characteristics, worker characteristics and firm characteristics listed in table 1. I report summary statistics for these variables for the population of surveyed applicants, and for the sub-sample of applicants whose cases end by a full tribunal hearing. Note that I include in particular two dummy variables allowing me to distinguish economic dismissals or redundancy payment claims from other cases, which is crucial as one may fear that the effect of economic conditions, if any, only concerns this type of cases. In table 2, I report the same summary statistics for the sub-sample of cases for which we know whether the worker was still unemployed at the time of the survey. Indeed, while all the survey variables I use come from the employers' responses, the employment status question is only asked to the dismissed employee. Given that there are moreover some missing responses to the employment status question, the sample for which the employment status is available is much smaller. However, for reasons that will become clearer in the next section, exploiting the information on workers' employment status is a crucial aspect of this work

All variables in $X$ are potentially correlated with case quality, but two among these variables are most likely to be a good measure of case quality. First, I define a dummy variable for bad misconduct: this dummy is equal to 1 if the reason for the workers' dismissal was misconduct in relation with health and safety (hygiene, smoking, drunkenness), violence or theft. This definition was chosen both on a priori grounds and because these "bad misconduct" cases have a significantly higher probability of being deemed fair dismissals by judges. Second, I use the settlement offer made by the firm to the worker: indeed, as the settlement offer is made by the firm to the worker in order to convince the latter to give up going to full tribunal hearing, it must be that the higher this offer given other characteristics, the more the worker is likely to prevail at trial, i.e. the higher the worker's case quality[7] (this argument is further developed in section 4.2). Note that the reason why settlement offers are lower for cases that go to full trial is because 80% of dismissed employees who do get an offer accept it, and therefore there is a high proportion (88%) of employees with no offers among those who go to full trial.

We use two variables to reflect economic conditions: the unemployment rate, which pertains to labour market conditions and therefore should affect workers relatively more than firms, and the bankruptcy rate, which should affect firms relatively more than workers. The unemployment rate we used is the claimant count rate in the region and month of application. Therefore, we have both cross-sectional (12 regions) and temporal variation. The bankruptcy rate is the yearly bankruptcy rate (VAT deregistration statistics, statistics available on the Small Business Service website, www.sbs.gov.uk) by industry and region; the identification comes from 3 years, 12 regions and 9 industries. As can be seen in table 1, the variation in economic conditions in the sample is quite substantial, so that prospects for meaningful estimation are good. Moreover, it is important to notice that the average unemployment rate and bankruptcy rate in the sample of applicants who go to full trial does not significantly differ from the average of these variables in the sample of all applicants (Tables 1 and 2). Therefore, it does not seem that the propensity of workers to go to full trial is correlated with economic conditions. Selection bias is thus unlikely to drive results on judges' decision as a function of economic conditions.

In the following section, I discuss how economic conditions can influence judges' decisions and how to estimate this effect empirically.

[7] In as much as firms anticipate that judges' decisions depend on economic conditions, controlling by settlement offers may dampen the direct effect of economic conditions on judges' decisions. Therefore, finding no effect of economic conditions on judges' decisions when controlling for this variable would not show that there is no effect, whereas finding some effect would consolidate the robustness of the results while indicating that firms may not have perfect information about judges' decision rule.

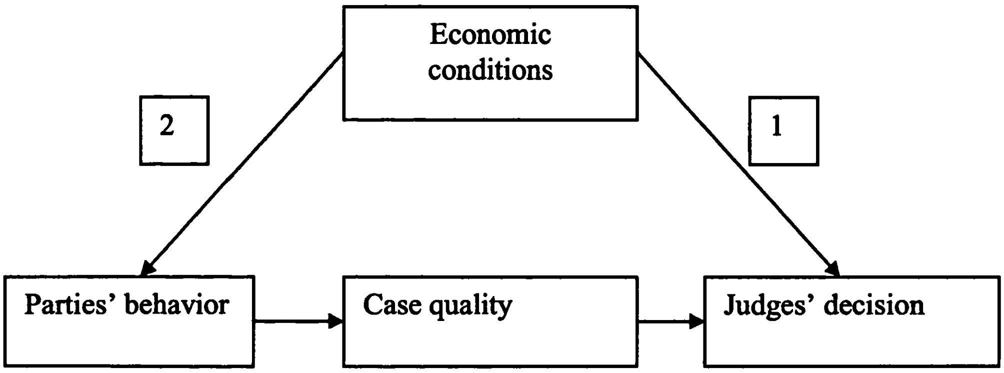
# 3 Models of judges' decision and the selection problem

Economic conditions can affect judges' decisions in two ways:

1.Directly, as an element taken into consideration in judges' decisions (channel 1 on figure 1).

2.Indirectly, by the influence they may have on the worker's and the firm's behaviour before the trial, affecting case quality (channel 2 on figure 1).

**Figure 1: the effect of economic conditions on judges' decisions**




## *3.1 The determinants of judges' decision making*

The reader is reminded that case quality refers to the quality of the worker's case, i.e. case quality is higher when the worker is more likely to prevail at trial. Judges' decision given case quality and economic conditions are independent of parties' behaviour. So, if case quality is perfectly observable, channel 2 can be ignored and one can directly analyze judges' decision as a function of case quality and economic conditions. Let $q$ be the case quality as perceived by the judges, and $u$ an indicator of economic conditions, such as the unemployment rate. Let $q^*$ be the judges' standard independently of economic conditions. Higher $q$ indicates better case quality and higher $u$ worse economic conditions. We can assume that the condition for the worker winning the case is:

$$q > q^* + u\alpha \qquad (1)$$

The right-hand side expression is the cut-off for the worker winning the trial: when this cut-off goes up, relatively higher quality cases end with a loss for the worker.

Hence, a higher right-hand side indicates that judges are more severe on workers. If $\alpha = 0$, then judges do not take into account economic conditions and their standard is $q^*$. If $\alpha > 0$, the cut-off goes up with worse economic conditions, i.e. judges are more severe on workers when economic conditions are worse. The opposite holds if $\alpha < 0$.

In the case where $\alpha \neq 0$, it is not obvious whether α should be positive or negative, i.e. whether, for a given $q$, judges should be more or less severe on workers when economic conditions are worse. Indeed, bad economic conditions have a negative impact on both firms and workers. They typically affect firms through lower profits and an increased bankruptcy risk, and workers through lower real wage growth and higher unemployment.

Judges can be assumed either to maximize welfare or to act strategically to please their constituencies, i.e. the workers and firms they represent. If judges try to maximize welfare, they can either try to maximize social welfare, or the welfare of the parties involved in each particular case. If judges try to maximize *social welfare*, they are confronted with the following trade-off. On the one hand, in bad times, financial pressure on firms increases, and so does the bankruptcy risk. Thus, any extra cost imposed on firms could have important consequences in terms of lost profits and lost jobs. On the other hand, as firing tends to be already high in bad times, being more severe on workers could encourage firms to fire even more, which would have adverse consequences for unemployment and aggregate demand. If the first effect dominates, then $\alpha > 0$, i.e. judges are more severe on workers in bad times compared to good times. If the second effect dominates, then $\alpha < 0$. If now judges try to maximize the *welfare of the parties*, they have to consider, in each particular case, whether the dismissed worker or the firm suffers more from degraded economic conditions. Relevant to this evaluation is the employment status of the plaintiff. Indeed, if the dismissed worker has already found a new job, worse or better economic conditions have little or no effect on his employment prospects. Therefore, for all cases where the worker is not unemployed, we expect, if anything, $\alpha > 0$, i.e. judges would favour firms when economic conditions are worse. If however the dismissed worker is unemployed, then, as in the case where judges try to maximize social welfare, the sign of $\alpha$ is undetermined; indeed if the worker is unemployed, clearly both the firm and the worker are likely to suffer from worse economic conditions. Thus, we can conclude that, if judges maximize the welfare of the parties, $\alpha$ is strictly higher if the worker is employed rather than unemployed at the time when his case reaches judgement. Another reason why we may expect the latter to be true is signalling: in the absence of perfect information, judges may take the employment status of the dismissed worker to be correlated with case quality, in the sense that if the employee is "good", and has indeed been "unfairly" dismissed, it is all other things equal easier for him to find a new job. When the unemployment rate is higher, it is however more likely that a worker is unemployed, which means that the bad signalling effect of being unemployed is attenuated. This signalling mechanism makes us expect that unemployed workers are less likely to win their cases in general, but relatively more likely to win their cases when economic conditions are worse.

Instead of trying to maximize welfare, lay judges may behave strategically and try to minimize their constituencies' dissatisfaction, and thus maximize their own popularity. Remember that a tribunal is composed of a chairperson, an employees' representative and an employers' representative. Clearly, firms as a group complain more about firing costs in bad times, hence firms' representatives are keener to please firms in bad times. On the other hand, as firings are more common in bad times, they can be perceived as a fact of life by workers as a group. Hence, higher firing in bad times would not be blamed so much on workers' representatives as on bad economic conditions. This can lead to firms' representatives exerting relatively more effort than workers' representatives to convince the chairperson in bad times. If this were not enough, firms' representatives could engage in intertemporal bargaining with workers' representatives and trade workers' victories in good times for firms' victories in bad times. This bargaining process is possible as lay judges typically work together on a series of cases[8]. Would the workers' representatives agree to this bargain? As long as firms' representatives' preference for more firms' victories in bad times relative to good times is stronger than workers' representatives' preference for more workers' victories in bad times relative to good times, the bargain is mutually beneficial. If so, lay judges would agree to be more severe on workers in bad times. This means that though the best cases would always win and the worst cases always lose, a case close

[8] In France, anecdotal evidence shows that such type of a bargaining is common in labour courts ("conseils de prud'hommes") : firms' representatives are usually small or medium businesses owners, so they trade-off big firms' victories for small firms' victories.

enough to the neutral judges' standard $q^*$ could lose in bad times and win in good times. The mechanism exposed above leads to $\alpha > 0$.

We can now summarize the expected effect of economic conditions on judges' decisions in table 3. The reader is reminded that $\alpha > 0$ means that judges tend to be more favourable to firms (and less to workers) when economic conditions are worse, and the opposite for $\alpha < 0$.

**Table 1: Theoretical predictions**

| | Judges' objective | | | |
|---|---|---|---|---|
| | Social welfare | Parties' welfare | | Judges' welfare |
| | | Worker employed | Worker unemployed | |
| **Sign of α** | $\alpha_{sw}$ ? | $\alpha_{pwe} > 0$ | $\alpha_{pwu} < \alpha_{pwe}$ | $\alpha_{jw} > 0$ |

Do we expect to see any differences in $\alpha$ depending on whether the bankruptcy rate or the unemployment rate is used to represent economic conditions? As argued above (2.2), the unemployment rate should affect relatively more the well-being of workers and the bankruptcy rate the well-being of firms. If judges aim at maximizing social welfare or their own welfare, this difference between the two variables does not have any obvious implications for $\alpha$. If judges aim at maximizing the parties' welfare however, this difference becomes relevant. If the worker is employed, any indicator of economic conditions is more likely to proxy for the conditions faced by the firm. On the other hand, the unemployed worker suffers more from a higher unemployment rate than from a higher bankruptcy rate, and hence we expect judges to be relatively more likely to decide in favour of the unemployed worker when the unemployment rate is higher rather than when the bankruptcy rate is higher. Given that we do not have a clear prior on whether the unemployment rate or the bankruptcy rate is a better indicator of firms' conditions as perceived by judges, we cannot draw a conclusion about which one should have a stronger effect in cases involving employed workers. But we can expect that using the unemployment rate as a measure of economic conditions should lead to judges being relatively more favourable to unemployed workers than using the bankruptcy rate. I.e. using the superscript *br* for the bankruptcy rate and *ur* for the unemployment rate, we expect that :

$$\alpha_{pwu}^{ur} < \alpha_{pwu}^{br}$$

## *3.2 The selection problem: parties' behaviour*

The former discussion assumes that case quality is perfectly observable. However, if we try to estimate $\alpha$ in (1) using data on cases that have reached the trial stage, we have to come to terms with the fact that case quality is imperfectly measured. Indeed, it can hardly be expected that a dataset collected by interviewing employers and employees, as detailed as it can be, should capture perfectly judges' view of case quality.

Assume then that for each case $i$ in the *population of applicants* to Employment Tribunals, the quality $q_i$ is given by:

$$q_i = X_i\beta_1 + \varepsilon_{1i} \tag{2}$$

where $X_i$ is a vector of observed characteristics for case $i$ and $\varepsilon_{1i}$ is a random error, normally distributed with zero mean. $X_i$ includes the constant and the control variables whose summary statistics are provided in table 1, section 2.2.

Moreover, assume that the judges' threshold is given by:

$$q^* = X_i\beta_2 + u\alpha + \varepsilon_{2i} \tag{3}$$

where $\varepsilon_{2i}$ is a random error, normally distributed with zero mean. The decision threshold is thus modelled in the same way as the case quality itself, i.e. assuming that the observer has noisy but unbiased information about its determination.

Then the empirical counterpart of equation 1 is a probit model. Thus, if win is a dummy variable taking the value 1 if the worker wins the trial and 0 otherwise, we have:

$$\begin{aligned} P(win_i = 1) &= P(q_i > q^* + u\alpha) \\ &= P(\varepsilon_{2i} - \varepsilon_{1i} < X_i(\beta_1 - \beta_2) - u\alpha) \end{aligned} \tag{4}$$

Crucially, one should note that the variable $win$ has missing values for all cases that do not reach the trial stage, i.e. the value of $win$ is observed conditional on the case reaching trial. Let $\hat{\alpha}$ be the estimate of $\alpha$ obtained by a probit estimation of equation (4) on the cases for which $win$ is observed.

Now, let I be an indicator variable taking the value 1 if a case reaches the trial stage and 0 otherwise. Suppose that applicants choose to go to trial if their case quality $q_i$

exceeds a certain threshold $q^I$ which depends on $X_i$ and possibly on $u$[9], so that $q^I = X_i\gamma_1 - u\delta + \varepsilon_{3i}$, where $\varepsilon_{3i}$ is a random error normally distributed with zero mean. Then the model for sample selection is given by:

$$P(I_i = 1) = P(q_i > q^I) = P(\varepsilon_{3i} - \varepsilon_{1i} < X_i\gamma + u\delta) \tag{5}$$

where $\gamma = \beta_1 - \gamma_1$.

Finally, the relevant model *for the selected sample* is:

$$P(win_i = 1 \mid I_i = 1) = P(X_i\beta - q^* - u\alpha + \varepsilon_{1i} - \varepsilon_{2i} > 0 \mid X_i\gamma + u\delta + \varepsilon_{1i} - \varepsilon_{3i} > 0) \tag{6}$$

Under those assumptions, two situations may arise:

- $\varepsilon_{1i} - \varepsilon_{2i}$ is uncorrelated with $\varepsilon_{1i} - \varepsilon_{3i}$: then $P(win_i = 1 \mid I_i = 1) = P(win_i = 1) = P(\varepsilon_{2i} - \varepsilon_{1i} < X_i(\beta_1 - \beta_2) - u\alpha)$. Thus $\hat{\alpha}$, the estimate of $\alpha$ obtained by using the probit model in (4) on the selected sample of cases reaching trial, does not suffer from any bias due to sample selection. Under this assumption, the fact that we do not perfectly observe case quality does not imply that we need to explicitly model the behaviour of parties before trial (channel 2 in figure 1) in order to get a consistent estimate of the effect of economic conditions on judges' decisions.
- $\varepsilon_{1i} - \varepsilon_{2i}$ is correlated with $\varepsilon_{1i} - \varepsilon_{3i}$: $\hat{\alpha}$ is then potentially biased. The two errors are likely to be correlated among themselves because they both include an omitted variable, the unobserved case quality $\varepsilon_{1i}$: for example, we may expect that all other things equal, cases with higher unobserved quality have a higher probability of reaching trial. Even so, if $\delta = 0$, i.e. economic conditions do not influence the selection process, $\hat{\alpha}$ will likely not be biased due to sample selection because the conditional mean of $\varepsilon_{1i} - \varepsilon_{2i}$ depends only on $X_i$ and not on $u$. If $\delta = 0$, ignoring the selection process and running the probit model (4) is equivalent to omitting a function of $X_i$. Given that $X_i$ is already included, this omission will bias the $X_i$ parameters but is unlikely to have much impact on the $u$ coefficient. However, looking at channel 2 on figure 1, one may think of a series of reasons why economic conditions can

[9] This assumption will be further justified by a model of parties' behaviour developed in section 5.1.

affect parties' behaviour before trial, leading to $\delta \neq 0$. For example, if the prospects on the labour market are bleak, the opportunity cost for a worker to go to trial may be lower, and therefore worse cases may proceed to trial. Under these circumstances, a higher $u$ leads to more cases with low unobserved quality being selected for trial, which then leads to a lower $\hat{\alpha}$. Thus, in general, if $\varepsilon_{1i} - \varepsilon_{2i}$ is correlated with $\varepsilon_{1i} - \varepsilon_{3i}$ *and* $\delta \neq 0$, $\hat{\alpha}$ captures the net effect of economic conditions on *both* parties' behaviour (channel 2) and judges' decisions (channel 1).

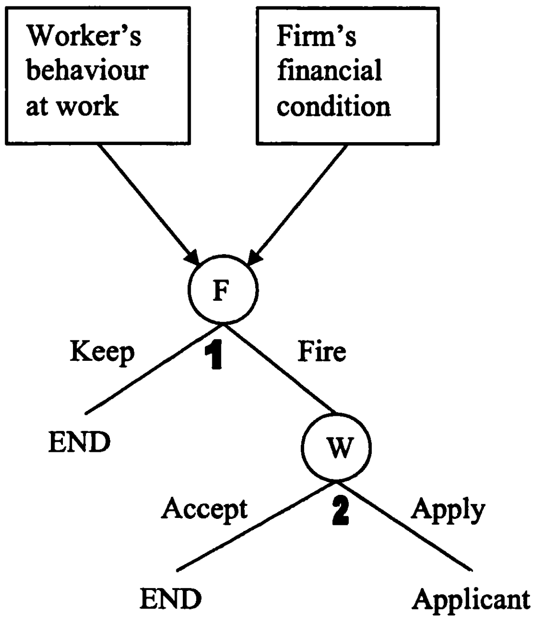
Therefore, determining the correct empirical strategy for estimating $\alpha$ in (1) requires examining the behaviour of parties before trial. However, before we can have a closer look at the determination of the selection process and hence $I_i$, we have to deal with a potential caveat. The whole discussion so far only takes into account the behaviour of parties from application to trial, whereas the behaviour of parties before application may also defeat the identification strategy. Indeed, the strategy strongly relies on the assumption, embodied in equation 2, that the case quality of applicants does not depend on economic conditions, i.e. $\varepsilon_{1i}$ is uncorrelated with *u*. Hence, we first have to ascertain whether such an assumption is reasonable, given that we do not observe $X_i$ for any case in which the employee does not apply to Employment Tribunal. This will be the purpose of the next section. We will then proceed to consider selection of cases *within* our sample in section 5.

# 4 The selection of the sample of applicants to Employment Tribunals

To deal with the selection of the sample of applicants, we investigate the typical process a case goes through before the application stage (Figure 2) in the United Kingdom. Each circle determines a decision point for an agent, F being the firm and W the worker. First, the firm decides whether to keep or fire the worker (node 1). Then, if the worker is fired, he decides to apply to the Employment Tribunal or not (node 2).

Given that I do not have information on $X_i$ for non-applicants, I cannot distinguish observed from unobserved case quality. I assume that if economic conditions affect case quality in some direction, then they will affect unobserved case quality in the same direction[10]. Thus, I examine if economic conditions affect the distribution of case quality among applicants in order to assess whether $\varepsilon_{1i}$, the unobserved component of case quality, is correlated with $u$ due to the selection of the sample of applicants.

**Figure 2: the selection process for applicants**




## *4.1 The effect of u on selection of applicants: theory*

Let $q_i$ be the case quality of an employed worker and $f_u(q)$ be the density of case quality among employed workers. This density may depend on $u$, because when unemployment is higher, employed workers are likely to shirk less in order to avoid getting unemployed when the value of unemployment is low.

Obviously, the lower the quality of the worker's potential case, the more likely the firm is to fire him. Then the worker is fired if:

$$q_i < q^F(u) \tag{7}$$

[10] This is not warranted under any possible set of hypotheses about the correlation matrix between unobserved case quality, observed case quality and economic conditions, but seems generally reasonable.

where $q^F(u)$ is the case quality threshold below which a worker is fired; this threshold may depend on $u$. Indeed, one may think that firms are likely to fire relatively higher quality workers when economic conditions are bad.
Then, among fired workers, all other things equal, workers with better case quality are more likely to apply. The worker applies to Employment Tribunal if:

$$q_i > q^A(u) \tag{8}$$

where $q^A(u)$ is the case quality threshold above which a fired worker decides to apply to tribunal; again this threshold may depend on $u$, as discussed in section 3.2.
Then the expected case quality of applicants is:

$$E(q \mid apply) = \int_{q^A(u)}^{q^F(u)} q.f_u(q)dq \tag{9}$$

With this notation in mind, let's discuss decisions at nodes 1 and 2.
At *node 1*, the population at risk is the entire population of employed workers. At this node, the firm decides to fire or keep the worker. As workers tend to shirk less in bad times, $f_u(q)$ the distribution of case quality in the population of employed workers presumably has a less thick lower tail: if $q^F(u)$ and $q^A(u)$ do not change, this effect is likely to increase $E(q \mid apply)$, the expected case quality of applicants. However, the decision rule of the firm itself is likely to change with economic conditions: indeed, in bad times, firms are less willing to keep relatively low productivity workers, so the firm will tend to fire workers with relatively higher case quality, i.e. $q^F(u)$ increases with u. Assuming no change in $q^A(u)$, $E(q \mid apply)$ is thus likely to increase when economic conditions are worse.
If the firm decides to fire the worker, at *node 2* the worker can accept the decision or apply to the Employment Tribunal. However, as we will discuss in more detail later on (section 5.1), it is not clear whether $q^A(u)$ increases or decreases with $u$.
Assuming that $f_u(q)$ and $q^F(u)$ change with $u$ as described above, we can conclude that if $q^A(u)$ is unaffected by $u$ or increases with $u$, then $E(q \mid apply)$ goes up when $u$ goes up. If, on the other hand, $q^A(u)$ decreases with $u$, the total effect of $u$ on $E(q \mid apply)$ depends on the relative magnitudes of the effects on $f_u(q)$ and $q^F(u)$,

which tend to increase $E(q \mid apply)$, and on $q^A(u)$, which tends to decrease $E(q \mid apply)$.

Given these possible correlations between case quality of applicants and economic conditions, in any subsequent regression analysis on the sample of applicants, the error, which possibly includes unobserved case quality, may be correlated with $u$. If unobserved case quality is indeed correlated with $u$ due to the selection of the sample of applicants, estimates of the effect of economic conditions on judges' decisions will be biased, even when controlling for selection within the sample of applicants. We therefore need to design an empirical strategy to estimate the effect of $u$ on $E(q \mid apply)$. If we can show that overall $u$ does not affect $E(q \mid apply)$, we may concentrate on selection within the sample of applicants.

## 4.2 Empirical strategy

The purpose of this section is to derive an empirical strategy to determine the correlation between the expected case quality of applicants, $E(q \mid apply)$, and economic conditions in the sample of applicants.

Observations on the total number of applications to Employment Tribunals can shed light on this issue. Indeed, looking at equation (9), we can derive a relationship between the number of applicants and the mean quality of applicants. Indeed, the number of applicants is proportional to:

$$P(q^A(u) \leq q \leq q^F(u)) = \int_{q^A(u)}^{q^F(u)} f_u(q)dq \qquad (10)$$

First, suppose employed workers do not react to economic conditions by exerting more or less effort to improve their potential case quality, so that $f_u(q)$ is unaffected by economic conditions. If so, it is obvious that if $q^F(u)$ increases with $u$ *and* $q^A(u)$ weakly decreases with u, then $P(q^A(u) \leq q \leq q^F(u))$, and hence the number of applicants, will increase. Then, if $f_u(q)$ does not change with economic conditions, $P(q^A(u) \leq q \leq q^F(u))$ can *only* weakly decrease with $u$ if $q^A(u)$ sufficiently increases with $u$. Therefore, if we find that the number of applicants does not increase with $u$, we can conclude that it is likely that $q^A(u)$ increases with $u$, i.e. workers are less willing to apply to Employment Tribunal when $u$ is higher. This in turn implies that if the number of applicants does not increase with *u,* then the mean case quality of applicants should be, if anything, higher.

If now employees react to worse economic conditions by shirking less, then $f_u(q)$ is affected by economic conditions and all other things equal the number of people below $q^F(u)$ decreases and so will possibly decrease the number of applicants. This effect in itself increases the case quality of applicants. However, if we take this effect into account, the conclusion that if the number of applicants does not increase with *u*, then the mean case quality of applicants increases is not as solid anymore. Indeed, now it could be that the number of applicants does not increase with *u although* $q^A(u)$ decreases with *u*. If $q^A(u)$ decreases with *u*, then workers, once fired, are more prone to apply, which in itself increases the number of applications and decreases the quality of applicants. But on the other hand, because workers shirk less, they are less likely to be picked upon in the first place, which diminishes the number of applicants and can compensate the positive effect of a decrease in $q^A(u)$ on the number of applicants. However, it remains true that if the number of applicants does not increase with *u*, then the mean case quality of applicants is more likely to increase than to decrease with *u*: indeed, any move in $q^A(u)$ would have to be big enough to compensate for the fact that less shirking means that $f_u(q^A(u))$ is smaller.

In general, we conclude that if the number of applicants does not increase with *u*, then the mean case quality of applicants is likely to weakly increase.

A second insight into the correlation between case quality of applicants and economic conditions is available using the micro dataset. Once the dismissed worker applies to the Employment Tribunal, the firm can offer an amount of money to the worker in order to settle the case instead of going to trial. It is reasonable to assume that the amount of the offer is, roughly speaking, proportional to the expected gains of the worker at trial, i.e. the probability of the worker winning multiplied by the monetary award he would get[11]. Thus, the ratio of the settlement offer B to the award A is a very good proxy for the probability of the worker winning according to the firm. Given A and B, we can therefore investigate the distribution of case quality among

[11] We will discuss more thoroughly a model of settlement behaviour in section 5.

applicants[12]. The micro dataset fortunately contains the amounts firms proposed to workers for a settlement and B is therefore known[13]. The awards workers would get if they won at trial are determined by the law and are a function of tenure, wage and age[14]; I can compute these amounts using the dataset and get A.

I can thus estimate the distribution of case quality in the whole sample using a kernel representation: I plot and compare the distribution of B/A in high unemployment versus low unemployment conditions, and high bankruptcy versus low bankruptcy conditions. If there is no difference in the distribution of B/A in low versus high unemployment conditions, we can conclude that the distribution of case quality of applicants is unlikely to be affected by economic conditions.

## 4.3 Results for the selection of the sample of applicants

The first test for selection bias is to examine the relationship between the number of applications to Employment tribunals and the unemployment rate. Burgess, Propper and Wilson (2001) find that there is none. Therefore, using the reasoning outlined in section 4.2 above, we conclude that the case quality of applicants is likely to weakly increase with worse economic conditions.

The second test for selection bias uses the distribution of firm's settlement offers as a proxy for the distribution of case quality. In the figure 3 below, we plot separately the distributions of case quality for high and low unemployment. As we can easily see, they are almost identical. As settlement offers are concentrated at 0, we may want to plot the settlement offers conditional on their being greater than 0 (figure 4). Again, the distributions for high versus low unemployment are essentially the same.

We proceed to do the same analysis for our second measure of economic conditions, namely the bankruptcy rate. We thus plot the distribution of case quality in low versus high bankruptcy conditions, for all cases (figure 5) and for cases with positive offers

[12] We do not have to assume here that the firm is perfectly informed. It is enough that the firm makes unbiased estimates of the workers' probability of winning at trial.

[13] In a certain number of cases, we only observe B if the offer was indeed accepted by the worker. Treating these cases separately in the analysis does not change the main results; hence, for simplicity, we ignore this distinction.

[14] The basic award is calculated by adding up the following amounts, but only continuous employment within the last 20 years can count: one and a half weeks' pay for each complete year of employment when an employee was between the ages of 41 and 65 inclusive; one week's pay for each complete year of employment when an employee was between the ages of 22 and 40 inclusive; half a weeks' pay for each complete year of employment when an employee was below the age of 22. As it happens, the basic award can be reduced or increased by the judge due to the specificities of each case. In fact, the award is almost never reduced, but rather increased. Thus, the basic award represents a good lower bound approximation for what the worker would get if he won at trial.

(figure 6). Although the distributions in high versus low bankruptcy rate are not as close to identical as in the case of the unemployment rate, they are still very similar so that it cannot be concluded that there is any significant difference, be it positive or negative.

As a further robustness check, I regressed the firm's settlement offer as a share of the workers' legally determined award on unemployment rate, bankruptcy rate, and the set of control variables. The results (not reported here) confirm the graphical analysis, showing no significant effect of either the unemployment or bankruptcy rate on case quality.

In conclusion, the tests performed are consistent with the hypothesis that case quality *of applicants* does not depend on the unemployment rate or the bankruptcy rate. We can therefore now concentrate on the selection of cases for trial *within* the sample of applicants.

# 5 The selection of applicants' cases to trial

Having established that the available empirical evidence is consistent with the absence of a correlation between case quality and economic conditions in the sample of applicants to Employment Tribunals, we can now concentrate on modelling the selection process of cases from application to trial. Modelling this process will ultimately allow us to give a behavioural basis to the selection equation (5).

## 5.1 A model of the selection of cases for trial

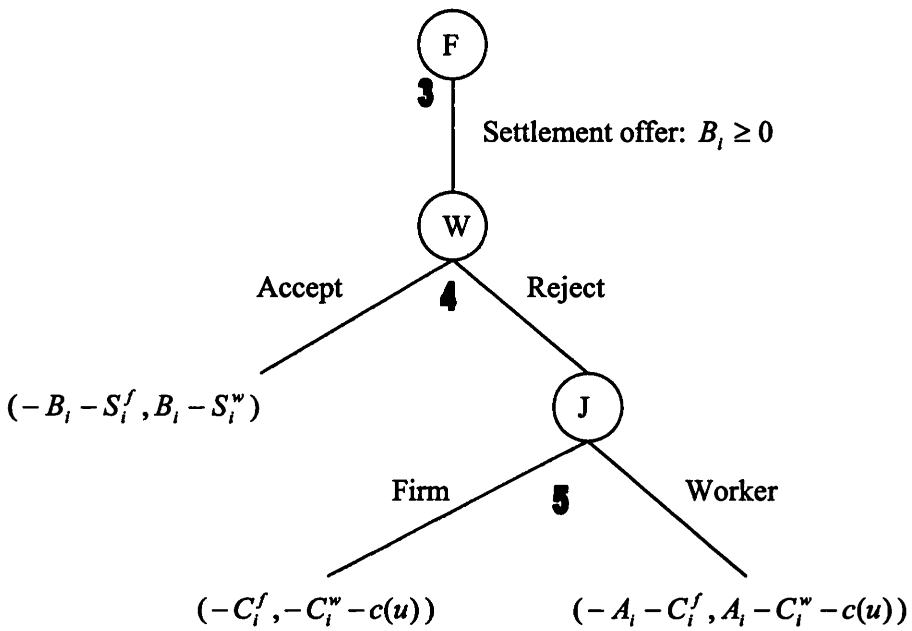
The only paper investigating the same question as ours, i.e. Ichino et al. (2004), uses a divergent expectations framework inspired by Priest and Klein (1984) to model the selection of cases for trial; to this divergent expectations framework, they add asymmetric stakes. Thus, a trial occurs for two possible reasons. First, a trial can occur because of divergent expectations. In this case, the worker and the firm disagree about the quality of the worker's case, the worker thinking his case is better than the firm thinks, and a trial occurs if the extent of the disagreement is big enough to make parties willing to incur the costs of a trial instead of agreeing on a settlement award, i.e. agreeing on how much the firm should pay the worker in order for the latter to drop his case. Second, a trial can occur because of asymmetric stakes, i.e. if the worker gains more than the firms loses from a trial. The resulting model predicts a lower quality of cases when unemployment is higher: this is because the alternative for the fired employee is either accepting dismissal and looking for a job in the labour market, or incurring trial costs and, if he/she wins, being reintegrated in his/her former job. The value of reintegration being higher in a depressed labour market, the cut-off for going to trial is lower: workers with less strong cases litigate when unemployment is higher, i.e. there is a negative selection bias. Because Ichino et al. find that, empirically, workers dismissed in a high unemployment context litigate more, have a lower case quality and at the same time win more often, they conclude that judges have a pro-worker bias.

The model designed by Ichino et al. is not applicable as such to the British case. In the United Kingdom, victory at trial is in practice almost never followed by reintegration, because losing firms are not forced to take back the victorious ex-employee. Instead, a financial compensation is awarded to the dismissed worker if the firm is found to have behaved unreasonably. As the financial compensation is set by a legal formula and does not depend on the unemployment rate, the worker does not gain more by going to trial in a high unemployment context. Hence, a negative selection bias is unlikely. But we cannot rule out the possibility that, for example, the time cost of trial is lower when unemployment is higher because job search is less efficient, which would also induce a negative selection bias.

To deal with this potential problem, we pursue our investigation of the typical process a case goes through before reaching trial in the United Kingdom: thus, figure 7 illustrates the decisions taken by parties from application to trial. As in figure 2, each circle determines a decision point for an agent, F being the firm, W the worker, and J the judges. If the worker applies, the firm decides on the amount of the settlement award it wishes to offer (node 3). Finally, if the worker rejects the firm's offer at node 4, the case proceeds to trial (node 5). These decisions will influence the distribution of quality among the cases reaching trial and will be the basis for the selection equation (5).

I now discuss the likely effect of economic conditions on decisions taken at each node in figure 7, looking at the quality of cases that proceed towards trial. I assume, as in section 4, that whichever effect economic conditions have on case quality at each decision node, the effect on unobserved case quality goes in the same direction, or is null.

**Figure 3: parties' behaviour before trial**




I build a model of the selection process of cases in the United Kingdom to determine how economic conditions can affect unobserved case quality through parties' decisions before trial.

The assumptions of the model are, as in Ichino et al.(2004), divergent expectations, to which we add an element of asymmetric information. The basic idea is that workers and firms start off with different beliefs about case quality because they have different information. The actions of each one of them act as signals and allow the other to update his beliefs. Economic conditions do not alter the information each party gets about case quality, but rather affect the decisions that are made based on this information. In other words, if economic conditions modify the pay-offs associated with different decisions, they modify the optimal decisions taken by agents and therefore the distribution of case quality for cases reaching trial.

First, we have to define the parties' beliefs about case quality. Assume the beliefs can be represented by probability distributions, in the Bayesian style. The belief of the worker involved in case $i$ is then represented by a random variable $Q_i^w$ and the belief of the firm involved in case $i$ is represented by a random variable $Q_i^f$; because the beliefs of the parties are about the same quantity, i.e. case quality $q_i$, $Q_i^w$ and $Q_i^f$ are positively correlated. Assume, moreover, that the best subjective estimate of the value of a variable about which the individual holds such a probabilistic belief, i.e. $Q_i^w$, is the expectation of that belief, i.e. $E(Q_i^w)$. If the belief has a normal distribution, this amounts, not surprisingly, to assuming that the best estimate is the *mean* of the distribution. An important assumption simplifying further reasoning is that the probability distributions of the beliefs of all firms on the one side, and of all workers on the other side, have the same shape and scale, and only differ in location. Intuitively, this means that all workers on the one side, and all firms on the other side, have the same degree of uncertainty in their beliefs, the only variation in beliefs coming from $E(Q_i^w)$. Thus, assuming, as seems reasonable, that beliefs reflect case quality, a worker with higher case quality has, on average, a higher $E(Q_i^w)$ but is not more or less certain of his case quality than a worker with a lower case quality.

Because each party updates her beliefs to incorporate what she learns from the other party's behaviour, we need to define beliefs about beliefs. Thus, $Q_i^{wf}$ is the firm's belief about the worker's belief, $Q_i^{wfw}$ is the worker's belief about the firm's belief about the worker's belief, etc.

In this framework, subjective probabilities of the worker's winning can be defined as follows:

$$P_i^w = P(Q_i^w > q^* + \alpha u) \tag{11}$$

$$P_i^f = P(Q_i^f > q^* + \alpha u) \tag{12}$$

In the same vein, beliefs about beliefs generate corresponding probabilities of the worker's winning, such as for example: $P_i^{wf} = P(Q_i^{wf} > q^* + \alpha u)$

The firm makes an offer $B_i \geq 0$ to the applicant. If the applicant accepts this offer, the parties' payoffs are:

$$U_i^{fS} = -B_i - S_i^f \tag{13}$$

$$U_i^{wS} = B_i - S_i^w \tag{14}$$

where $S_f$ and $S_w$ are settlement costs, and the superscript S stands for settlement.

Assuming that the parties are risk neutral, we can define their expected utilities if they go to trial as follows:

$$U_i^{fT} = -P_i^f A_i - C_i^f \tag{15}$$
$$U_i^{wT} = P_i^{w2} A_i - C_i^w - c(u) \tag{16}$$

where the superscript T stands for trial, $P_i^{w2}$ is the belief of the worker about his probability of winning given the offer $B_i$, $C_i^f$ and $C_i^w$ are litigation costs, for the firm and the worker respectively and $A_i$ is the size of the stake, or award the worker would get if he won. *c(u)* is a cost or benefit incurred by the worker if he litigates, and it is assumed to be a function of economic conditions. Indeed, at first, the dismissed worker is unemployed, and bad economic conditions render job search less efficient. If searching for a job and taking care of an Employment Tribunal case are alternative uses of time and money, then a change in the returns to job search will affect the decision to invest in an Employment Tribunal case. However, the effect of economic conditions on the latter decision is ambiguous. Indeed, on the one side, a lower return to job search would all other things equal encourage unemployed workers to pursue their cases. But if the negative impact of litigation on the prospects of finding a job is considerably amplified by worse economic conditions, then this would incite unemployed workers to litigate less. To simplify, one can assume that once the dismissed worker finds a new job, economic conditions do not affect him any more. So, if the worker is employed when he decides whether to settle or to take the case to full tribunal hearing, then the current economic conditions have no effect on his decision. To summarize, then, if the worker is not unemployed, we can assume that *c(u)=0*, so that economic conditions have no effect on the selection of cases for trial.

How is $P_i^{w2}$, the updated belief of the worker about his probability of winning defined? First, note that for the firm to make an offer $B_i \geq 0$ to the applicant, it must be true that the value of a settlement for the firm is higher than the value of going to trial:

$$U_i^{fS} \geq U_i^{fT} \Leftrightarrow B_i \leq P_i^f A_i + C_i^f - S_i^f \Leftrightarrow P_i^f \geq \frac{B_i - C_i^f + S_i^f}{A_i} \tag{17}$$

Therefore, relying on the above observation, the updated probability of the worker's winning according to the worker is:

$$P_i^{w2} = P\left( Q_i^w > q^* + \alpha u \mid P_i^f \geq \frac{B_i - C_i^f + S_i^f}{A_i} \right) \tag{18}$$

Let $Q_i^{w2}$ be the updated belief of the worker about the quality of his case. $Q_i^{w2}$ is defined by its cumulative distribution function:

$$F_{Q_i^{w2}}(x) = P\left(Q_i^w < x \mid P_i^{fw} \geq \frac{B_i - C_i^f + S_i^f}{A_i}\right) = P(Q_i^{w2} < x) \tag{19}$$

The formation of $Q_i^{w2}$ sheds some light on the reason why many firms choose to make an offer equal to zero. Indeed, a zero offer tells the worker that his case quality is below a certain threshold, whereas a positive offer tells him that his case quality is above the threshold. When case quality is low, the expected value of case quality above the threshold is much higher than the case quality. Therefore, when the worker's case quality is relatively low, any positive offer will result on average in the worker updating his belief upwards to a considerable degree, which in many cases will lead him to decline the firm's offer and go to trial. Anticipating this, the firm does not make any positive offer in the first place.

In general, the worker decides to reject the firm's offer and go to trial if:

$$\begin{aligned} U_i^{wS} < U_i^{wT} &\Leftrightarrow B_i - S_i^w < P_i^{w2} A_i - c(u) - C_i^w \Leftrightarrow P_i^{w2} > \frac{B_i - S_i^w + c(u) + C_i^w}{A_i} \\ &\Leftrightarrow P(Q_i^{w2} > q^* + \alpha u) > \frac{B_i - S_i^w + c(u) + C_i^w}{A_i} \end{aligned} \tag{20}$$

Remember that probability distributions of $Q_i^{w2}$ only differ among workers by their location. We can hence define a function $h$ such that condition (20) for going to trial can be rewritten as:

$$E(Q_i^{w2}) > h\left(\frac{B_i - S_w + c(u) + C_w}{A_i}, q^* + \alpha u\right) \tag{21}$$

The above condition gives behavioural foundations to the selection equation 5: it says that if the plaintiff's best estimate of case quality is above a certain threshold[15], then the plaintiff proceeds to trial. As $h$ is increasing in its two arguments, we can derive the effect of an increase in any of the variables.

We have:

$$\frac{\partial h}{\partial B_i} > 0$$

[15] If we did not assume that all workers have beliefs with the same shape and scale, this threshold would also depend on the distributional form of each worker's belief, and not only on the specified variables.

Indeed, the higher the offer $B_i$, the less workers are willing to go to trial.

Then:

$$\frac{\partial h}{\partial A_i} < 0$$

because higher awards make workers more willing to go to trial.

And:

$$\frac{\partial h}{\partial C_i^w} > 0$$

because as trial gets more costly, workers are less willing to go to trial.

Symmetrically, we have:

$$\frac{\partial h}{\partial S_i^w} < 0$$

because as settlement gets more costly, it is relatively less costly to go to trial.

The effect of *u* on *h* is ambiguous, as it depends on the unknown function *c(u)* and $\alpha$.

We note that *c(u)* plays the same role as $C_i^w$, so:

$$\frac{\partial h}{\partial c(u)} > 0$$

Hence, for a given $\alpha$, if *c(u)* goes up with *u*, then less bad cases go to trial, and conversely. However, as argued above, if the worker is employed then we assume that *c(u)=0.*

As for $\alpha$, we have:

$$\frac{\partial h}{\partial \alpha} > 0$$

So, if $\alpha > 0$, then *h* goes up with *u*, and conversely. This means that if judges are more severe with workers when economic conditions are worse, then workers are more likely to settle instead of going to trial as economic conditions deteriorate, and the opposite holds if judges are less severe with workers when economic conditions are worse.

We are now ready to proceed to the empirical specification.

## 5.2 Empirical specification

### 5.2.1 The selection equation

Using condition (21), we can derive a probit model for a case going to trial. Define trial as a dummy variable taking the value 1 if the worker goes to trial and 0 otherwise. To derive the empirical counterpart of the condition for going to trial, we must specify the worker's estimate of case quality $E(Q_i^{w2})$ and the $h$ threshold as a function of observed variables.

Assuming that $E(Q_i^{w2})$ is unbiased, we can define $E(Q_i^{w2})$ as:

$$E(Q_i^{w2}) = q_i + \varepsilon_{4i} = X_i\beta_1 + \varepsilon_{1i} + \varepsilon_{4i} \tag{22}$$

where $\varepsilon_{4i}$ is normally distributed with zero mean, and $\varepsilon_{1i}$ is the error defined in equation (2), i.e. the error associated with the distribution of case quality among applicants. The reader is reminded that the vector $X_i$ includes the offer $B_i$ made by the firm, so that the empirical specification is consistent with the definition of $Q_i^{w2}$ in (19).

Assume moreover that the costs $C_i^w$ and $S_i^w$ are defined by linear combinations of the variables in $X_i$. Given that we have also included the variables determining $A_i$ in $X_i$, we may now approximate $h$ as a linear function:

$$h = X_i\gamma_1 + u\delta + \varepsilon_{5i} \tag{23}$$

where $\varepsilon_{5i}$ is normally distributed with zero mean.

Therefore, the empirical counterpart to equation (21) is the probit model given in equation (5) which we can now reformulate as:

$$\begin{aligned} P(trial = 1) &= P(X(\beta_1 - \gamma_1) + u\delta + \varepsilon_{1i} + \varepsilon_{4i} - \varepsilon_{5i} > 0) \\ &= P(X_i\gamma + u\delta + \varepsilon_{1i} - \varepsilon_{3i} > 0) \end{aligned} \tag{24}$$

where $\varepsilon_{3i} = -\varepsilon_{4i} + \varepsilon_{5i}$. Thus, whereas in the formulation in section 3.2, equation (5), we explicitly included a single error term, $\varepsilon_{3i}$, for the selection of a case for trial, we have now shown that this error has two empirically undistinguishable components, $\varepsilon_{4i}$, the error in worker's belief about his case quality, and $\varepsilon_{5i}$, the error coming from our failure to perfectly observe the threshold $h$.

$\delta$ represents the effect of economic conditions on the decision to go to trial. Remember that if $\delta = 0$, then the correct estimation of $\alpha$ in equation 4 on the selected sample does not require an explicit modelling of the selection process.

### 5.2.2 The win equation

As argued in section 3.2, the correct specification for the win equation depends on assumptions about the correlation between unobserved case quality and economic conditions in the sub-sample of cases reaching trial.

First, we can make the very restrictive assumption that there is no effect of economic conditions on case quality (observed or unobserved) at trial. If so, we can use a macro time series of the percentage of cases reaching the trial stage that have been concluded with a worker victory and directly regress this variable on the time series of unemployment rates using ordinary least squares.

Second, we can relax the previous assumption and assume that while economic conditions may have an effect on dismissed workers' decision to go to trial, this effect is fully captured by observed variables *other* than *u*, so that $\delta = 0$. Then, the correct specification is given by equation (4), i.e.:

$$\begin{aligned} P(win_i = 1) &= P(q_i > q^* + u\alpha) \\ &= P(\varepsilon_{2i} - \varepsilon_{1i} < X_i(\beta_1 - \beta_2) - u\alpha) \end{aligned} \quad (4)$$

In all the probit specifications we use, standard errors are clustered by region as our main variable of interest, the unemployment rate, is taken at the regional level. Moreover, for cases tried in the same region, decisions may be taken by the same judges.

Remember that the economic conditions variables are defined by month and region for the unemployment rate, and by year, region and industry for the bankruptcy rate. One can thus ask to what extent cross-sectional versus temporal variation is important in explaining trial outcome: are workers from regions or industries facing a worse economic situation more likely to win/lose at trial, or is it the change in economic conditions over time that determines whether workers are more or less likely to prevail at trial? To answer this question, we run the probit specification with different sets of fixed effects: region effects, region and industry effects, and finally region, industry and year effects.

Third, we can further relax our assumptions, allowing for selection on unobservables. Thus, assuming that $\varepsilon_{1i} - \varepsilon_{2i}$ is correlated with $\varepsilon_{1i} - \varepsilon_{3i}$, we have to estimate equation (6):

$$P(win_i = 1 \mid I_i = 1) = P(X_i\beta - q^* - u\alpha + \varepsilon_{1i} - \varepsilon_{2i} > 0 \mid X_i\gamma + u\delta + \varepsilon_{1i} - \varepsilon_{3i} > 0) \qquad (6)$$

For this purpose, one can use a Heckman-style strategy (Heckman, 1979, Van de Ven and Van Praag, 1981), and a maximum likelihood technique. This technique has the advantage of giving an estimation of $\rho$ (rho) the correlation between $\varepsilon_{1i} - \varepsilon_{2i}$ and $\varepsilon_{1i} - \varepsilon_{3i}$. However, as argued by Sartori(2003), the estimator may perform poorly as the same variables are included in both the selection (trial) and outcome (win) equations. The mediocre performance of Heckman estimators is particularly problematic in small samples, and our sample is indeed relatively small, especially if we want to include the variable documenting the worker's employment status in the estimation. However, given the structure of our problem it seems reasonable to assume that $\varepsilon_{1i} - \varepsilon_{2i}$ is strongly correlated with $\varepsilon_{1i} - \varepsilon_{3i}$. Radicalizing this assumption to $\varepsilon_{1i} - \varepsilon_{2i} = \varepsilon_{1i} - \varepsilon_{3i}$, so that the unobserved component of case quality is the same in the decision of the worker to go to trial as in the decision of the judge, we can use the maximum likelihood Sartori estimator to derive an estimation of $\alpha$. Note that in both the Sartori and Heckman estimations, we do not include any region, industry or year fixed effects, as the maximum likelihood estimation algorithm does not converge if any fixed effects are included.

However, the Sartori estimator will only be more accurate than the simple probit if the hypothesis of identical errors is justified. In general, given the properties of the Sartori estimator, we know that it will provide an upper bound (in absolute value):

$$|\alpha_{probit}| < |\alpha_{true}| < |\alpha_{Sartori}| \qquad (25)$$

Without going into technical details, we can intuitively explain why the Sartori estimator is an upper bound in our framework. Indeed, assuming that the worker is aware that, say, $\alpha_{true}>0$, i.e. judges are more severe on workers when economic conditions are worse, the worker will be less willing to go to trial in bad times, and this will lead to relatively higher unobserved case quality at trial. Therefore $\alpha_{probit}$, not taking into account this selection bias, would underestimate the real effect of economic conditions. This implies that, by contrast, if there is no selection bias, $\alpha_{Sartori}$ will overestimate the effect of economic conditions.

The Sartori identifying hypothesis of identical errors may not be accurate in two cases. First, the hypothesis is not justified if our observed variables are an excellent measure of case quality so that there is no systematically unobserved case quality but mainly noise. Second, the hypothesis is flawed if the unobserved component of case quality according to the worker is largely uncorrelated with the unobserved component of case quality according to the judge. We have good reasons to believe that the correlation is less than one, as both firms and workers are likely to be surprised by judges' decisions. For example, 30% of firms and 64% of workers say they did not expect the outcome of the trial. Moreover, among cases reaching trial with the firm certain that it would win, 40% still end up with a worker victory!

### *5.2.3* Taking into account the employment status of the worker

We have argued that case quality is unlikely to be correlated with economic conditions for workers who were employed at all nodes where they had to take a decision. Therefore, selection bias due to worker behaviour before trial depends on the worker's employment status. Moreover, distinguishing between employed and unemployed workers allows us to test whether, assuming judges' objective is the parties' welfare, it is indeed the case that judges are relatively more lenient with unemployed workers when economic conditions are worse, i.e. whether $\alpha_{pwu} < \alpha_{pwe}$ (table 3).

As explained in section 2.2, we know whether the worker was unemployed or not at the time of the survey for a sub-sample of our data. The survey takes place shortly after the case is finished. Hence, we can reasonably hypothesise that if a worker is unemployed at the time of the survey, he was unemployed at all the moments when he had to take decisions (node 2 in Figure 2, and node 4 in Figure 7). Hence, observing the effect of economic conditions on the selection of unemployed workers for trial allows us to estimate $\beta_1$ in equation (24). Conversely, if a worker is employed at the time of the survey, we are not sure what his employment status was before the survey; however, as interviews generally take place shortly after the end of the case, it is likely that the worker had found a job by the time he/she reached the trial stage.

The number of unemployed workers in the sample is small (84 in the dataset, of which 35 reach the trial stage), hence estimating on the sample of unemployed workers alone is likely to lead to unreliable results. Moreover, I want to compare the effects of economic conditions on unemployed vs. employed workers. Therefore, I use a dummy for the employment status and interact economic conditions with this dummy. Thus, let $U$ be a dummy taking the value 1 if the worker is unemployed at the time of the survey, and 0 if the worker found a new job[16].

We re-run the probit, Heckman and Sartori regressions adding the $U$ dummy variable for being unemployed, and the interaction terms between the unemployment rate and $U$, and bankruptcy rate and $U$. This yields equation (26) for the probit estimator and equation 27 for the Heckman and Sartori estimators.

$$
\begin{aligned}
P(win_i = 1) &= P(q_i > q^* + u\alpha + U\varphi + uU\alpha_U + \varepsilon_{2i}) \\
&= P(\varepsilon_{2i} - \varepsilon_{ii} < X_i(\beta_1 - \beta_2) - u\alpha - U\varphi - uU\alpha_U)
\end{aligned}
\tag{26}
$$

$$
\begin{aligned}
P(win_i = 1 \mid I_i = 1) = P(&X_i\beta - q^* - u\alpha - U\varphi - uU\alpha_U + \varepsilon_{1i} - \varepsilon_{2i} > 0 \\
&\mid X_i\gamma + u\delta + U\delta_U + uU\delta_{Uu} + \varepsilon_{1i} - \varepsilon_{3i} > 0)
\end{aligned}
\tag{27}
$$

Note however that because we have a nonlinear model, the marginal effects are not rendered by the coefficient on the interaction term but must instead be computed separately. For probit models, the Stata programme inteff takes care of this calculation. However, for the Heckman and Sartori estimators, no such calculating module exists, which means that the coefficients on the interaction terms in the *outcome* equation ***should not*** be interpreted as marginal effects. For the *selection* equation (24) however, we can run separately a probit and calculate the marginal effects for the interacted terms; these marginal effects will be the same for Heckman and Sartori, because the selection equation is a probit in both cases, and if coefficient estimates may be slightly different, it is due to different numerical approximations.

We are now ready to examine the empirical results.

# 6 Empirical results

This section analyses the results stemming from the estimation of the models discussed in the previous section.

[16] This variable is constructed in such a way that inactive workers are excluded. There are only 4 inactive workers in the sample, and our prior about judges' attitude towards them is not clear-cut; therefore we concentrate on estimating the difference between employed and unemployed workers.

First, we assume that there is no effect of economic conditions on case quality (observed or unobserved) at trial, which allows us to use a macro time-series. The micro data we use only covers a period of two years. To get a broader picture, we plot the yearly win rate in unfair dismissal cases (from Burgess et ali., 2001) against the unemployment rate on the period 1985-2001 (Figure 8). The graph shows a negative relationship between the percentage of workers' victories and the unemployment rate, which is confirmed by the corresponding OLS regression. Thus, a one-point increase in the unemployment rate is significantly associated with a one-point decrease in the proportion of workers prevailing at trial, implying that $\alpha > 0$.

Assuming that the effect of economic conditions on case quality, if any, is captured by our control variables, we can directly retrieve $\alpha$ by estimating equations (4) (without control for the worker's employment status) and (26) (with control for the worker's employment status) by a probit model (Table 4). Columns 1 to 4 of table 4 estimate equation (4) with different fixed effects, whereas columns 5 to 8 estimate equation (26), again with different fixed effects.

Bear in mind that our favourite estimates for the effect of economic conditions on judges' decisions are to be found in columns 1-4, for reasons that will become clearer as we proceed.

Coefficients on control variables are reported in table A-1 of the annex and, for the sake of brevity and focus, they will only be partially discussed here.

The negative effect of worse economic conditions on workers' probability of prevailing at trial is consistent across all estimations in table 4. In *column 1*, where no fixed effects are added, the effect of being in a month and region with an unemployment rate higher by one point is to significantly diminish the probability of the workers' winning by 3.3 points. Similarly, the effect of being in an industry-region-year with a bankruptcy rate higher by one point is to decrease the worker's probability of winning by 1.6 points. Note that excluding the characteristics controls leads to similar point estimates, although slightly lower in absolute value for the unemployment rate (results not reported here). This suggests that the inclusion of individual characteristics does not have a big effect on the estimates of the effect of economic conditions. Therefore, assuming that there is no selection on unobservables, the results obtained on the macro series and reported on figure 8 above give a reasonable approximation for the effect of the unemployment rate on judges' decisions.

Moving to *column 2*, the addition of region effects has little impact on the coefficient on the bankruptcy rate, but, interestingly enough, it more than *doubles* the coefficient on the unemployment rate, implying that a worker applying to the Employment Tribunal in a month where the unemployment rate is higher by one point sees his probability of prevailing at trial diminish by 7.7 points. This result is important as one may have been worried ex ante about the fact that unobserved differences across regions drive the results. Instead, in both macro and micro data the *time* variation in unemployment does make a difference to Employment Tribunals outcomes.

Adding industry dummies in *column 3* does not affect the size of the coefficient on the unemployment rate, although significance is reduced, falling slightly below the 10% level. However, the inclusion of industry dummies doubles the coefficient on the bankruptcy rate: thus, a worker applying to the Employment Tribunal in a year where the bankruptcy rate is higher by one point sees a 2.7 point decrease in his probability of winning his case. This implies that for the bankruptcy rate as for the unemployment rate, time variation has larger effects than cross-sectional (region and industry) variation.

In *column 4* at last, we also include a year dummy to account for time variation. This does not have any dampening impact on the estimates of the effect of economic conditions: on the contrary, both coefficients are still significant and higher in absolute value, with the coefficient on the unemployment rate even doubling again.

In *columns 5 to 8*, the sample is reduced as we now want to control for the employment status of the worker. First, note that the effect of being unemployed on the worker's probability of winning is extremely strong and significant in all specifications: being unemployed makes the worker 64% to 82% more likely to lose his case[17]. This suggests that the case quality of unemployed people is lower. However, this may be for two reasons: either unemployed workers indeed come to trial with lower quality cases, one of the possible reasons being that, all other things equal, they litigate more (we will test this in the next table). Or judges *believe* that unemployed workers are worse. This occurs because of the signalling problem described in section 3.1: i.e. if judges are not perfectly informed, observing that a worker was unable to find a job signals to them that this worker is likely to be a

[17] If we perform a probit estimate without controls, we find that the marginal effect of being unemployed on the probability of winning is significantly negative at the 1% level and of similar magnitude (61%).

lemon, and that therefore the firm legitimately fired him. The observed coefficient on the interaction between the unemployed dummy and the unemployment rate is however positive (columns 5-8), indicating that unemployed workers have a higher probability of winning their cases when the unemployment rate is higher. This interaction effect, at first significant (column 5), becomes larger when adding fixed effects (columns 6-8), though is falls short of statistical significance in columns 6-7. It is worth noticing that if we perform the specification in column 1 on the set of unemployed workers only (28 observations available), thus allowing the coefficients on control variables to differ for the unemployed, we find a significant positive coefficient on the unemployment rate, the magnitude of the coefficient being very similar (0.058) to the one found in column 5. If instead we interact the unemployed dummy with the bankruptcy rate, results are weaker and often insignificant (results not reproduced here). This confirms our hypothesis (section 3.1) about the difference between the unemployment rate and the bankruptcy rate taken as measures of economic conditions when judges are maximizing the parties' welfare: as unemployed workers are more affected by the unemployment rate than by the bankruptcy rate, a change in the unemployment rate has a bigger and more significant effect for them than a change in the bankruptcy rate.

Let us now comment on the coefficients on the economic conditions indicators. Overall, the inclusion of different sets of fixed effects has an effect on the coefficients that is very similar to the one observed in columns 1-4[18]. Noticeably, coefficients on both economic conditions variables tend to be higher in columns 5-8 than in columns 1-4. However, this is not due to the explicit inclusion of the employment status variable, but rather to the sub-sample used: indeed, performing the regressions 5-8 on the same sample but excluding the employment status dummy and the interaction of the latter with the unemployment rate yields very similar estimates, except for column 1, where the estimate of the unemployment rate coefficient in the absence of control for unemployment status is lower (-0.032). Although, due to the small sample, some doubt about the precise magnitude of the coefficients in columns 5-8 is permitted, using the employment status of the worker has allowed us to confirm that, assuming

[18] The only noticeable difference is the very sizeable jump in the coefficient on the unemployment rate when moving from column 7 to column 8 (adding the year dummy), whereas the jump from column 3 to column 4 was less important (though still big). I explored this issue to find out that the main jump in the coefficient is due to including together region and year dummies. I do not have a good explanation why this jump should be so important and hypothesize that it is simply a random variation due to the small number of observations (93).

that there is no selection on unobservables, we have $\alpha_{pwu} < \alpha_{pwe}$ (table 4), i.e. worse economic conditions decrease more the probability of winning for employed relative to unemployed workers. Moreover, a new finding emerges: unemployed workers, while having a significantly lower probability of winning on average, are actually *favoured* by judges when the unemployment rate is higher, i.e. $\alpha_{pwu} < 0$ if the unemployment rate is used as the measure for economic conditions.

Now, in a third stage, we allow for economic conditions to influence the selection of cases for trial, even conditionally on observed variables (table 5). As discussed above, we use two estimation techniques, namely Heckman and Sartori. We note that in the selection equation (lower part of the table), the effect of economic conditions on a case being selected for trial is very close to 0 and insignificant in all columns. Therefore, unsurprisingly, the outcome equation of the Heckman estimator (columns 1 and 3) gives results that are almost identical to the probit models (columns 1 and 5 of table 4), although somewhat less significant. Not only is the effect of economic conditions unaffected by the inclusion of the selection equation, but the likelihood ratio test does not reject the hypothesis that the outcome equation and the selection equation are independent, i.e. that there is no selection on unobservables. However, as explained in section 5.2.2 above, we may be concerned about the fact that the Heckman estimator is inefficient due to the inclusion of the same variables in the selection and outcome equations. Therefore, as a robustness check, we use the Sartori estimator in columns 2 and 4, which also yields significant and negative coefficients on the economic conditions variables. In column 1, the Heckman estimation technique gives us an estimator of rho, 0.689, which happens to be closer to 1 than to 0; thus, the basic assumption of the Sarotri estimator, i.e. a rho equal to 1 appears to be a reasonable if somewhat poor approximation. As expected, the coefficients on economic conditions tend to be bigger in absolute value in column 2 than in column 1. In column 3, the best estimate of rho is negative, which explains why assuming a positive correlation equal to 1 yields this time lower estimates (in absolute value) on economic conditions in column 4 versus column 3. We will not delve more into comparing the Heckman and Sartori estimators as it appears that taking into account the effect of economic conditions on case selection does not significantly change the results given by the simple probit estimation. To close the discussion of table 5, we observe that unemployed workers are not more likely to go to trial than employed workers; the point estimate, though insignificant, even indicates that unemployed workers are less likely to go to trial[19]. Moreover, worse economic conditions do not make it more likely for unemployed workers to go to trial, and therefore, we cannot reject that $\beta_1 = 0$ in equation (24).

Before concluding this section, let us make a few comments on the effects of control variables reported in the annexes. First, the variables we thought proxy best for case quality do indeed yield consistent results: a higher settlement offer is generally associated with a higher probability of the worker winning, and the bad misconduct dummy always has a negative and significant effect on the worker's probability of winning. Second, we do distinguish dismissals for economic reasons, and we find that these cases usually lead to a lower probability of the worker winning the case[20]. Third, contrary to what the formulation of the law would make us expect (section 2.1), it does not seem that the size and administrative resources (personnel department) of the firm have a significant impact on trial outcomes, even when explicitly controlling for case selection. Fourth, consistent with the lesson from the trucker's example (section 2.1), the use by the firm of an internal procedure makes it more likely for the firm to prevail at trial. Finally, all other things equal, workers with higher wages in their lost job are more likely to lose at trial than workers with lower wages. We hypothesise that this is due to the fact that workers with higher wages would get higher awards if they were to win, and judges may be more demanding with cases implying higher payments from the firms to the workers, i.e. judges' threshold increases with the worker's (past) wage.

Given the above discussion, our favourite set of estimates is to be found in columns 1-4 of Table 4. Indeed, controlling for the employment status of the worker does not change the basic probit estimates of the effect of economic conditions on judges' decisions, but only forces us to work with a smaller sample; it is only interesting to control for the worker's employment status to determine the specific effect on unemployed workers, but not to compute the overall average effect. As for the

[19] One may ask whether unemployed workers are also those who apply to Employment Tribunals multiple times. First, most workers in the sample (97.7%) are bringing forward their first application ever, and the unemployed are if anything less likely to have brought an application before, the difference between the two groups being statistically insignificant.

[20] Note that excluding altogether economic dismissals from the whole analysis does not change the basic results.

selection models presented in table 5, they are rejected by the data, and should therefore better be seen as a robustness check.

# 7 Conclusion

This study has shown that economic conditions such as the unemployment rate and the bankruptcy rate affect the implementation of Employment Protection Legislation. In the United Kingdom, judges tend overall to decide more frequently in favour of firms when unemployment or bankruptcy rates are higher. However, judges' decision rule is different depending on the dismissed worker's employment status: the unemployment rate has a negative effect on the probability of dismissed workers who have found a new job winning their cases, whereas the effect for unemployed workers is positive.

Among the theories of judges' decision discussed in section 3.1, the empirical results mostly support the theory that judges' objective is to maximize the joint welfare of the parties involved in each case. The results do not rule out that judges also try to maximize social welfare or their chances of remaining in office, but make these objectives less likely. First, assuming that judges try to maximize welfare, the finding that unemployed workers get a different treatment is somewhat surprising. To understand this point, assume that the function relating optimal firing costs to economic conditions is monotonic. Given that judges in general are more pro-firm when economic conditions are worse, they may think that firing costs should be lower under such circumstances. But then it is not quite consistent for them to decide more often in favour of unemployed workers when unemployment is higher. Indeed, the higher the unemployment rate, the more likely it is that dismissed workers will remain unemployed. Therefore judges would tend to be less and less favourable to firms as economic conditions get worse, which would defeat their initial purpose.

Now, assuming that judges try to maximize their probability of remaining in office, the differential treatment of employed and unemployed workers does not seem to be justified either. Judges representing workers must have the support of workers' unions, and of firms' groups for judges representing firms. However, whereas it is reasonable to assume that these organizations have an idea about how often judges decide in favour of workers, it is difficult to believe that members of organizations who are not directly involved in the case would have any information about whether the worker was employed or not in that particular case. Even if they did have this information, it is not clear whether they would retain it as relevant. This means that it is difficult for outsider observers to spot the beneficial treatment that judges confer on unemployed workers in high unemployment contexts, and therefore such beneficial treatment should be of no or little interest to judges only concerned about maximizing their popularity. Therefore, I conclude that the observed behaviour of judges is mostly consistent with their maximizing the joint welfare of the parties involved in each case. However, judges' maximizing the joint welfare of the parties may generate a negative externality. Indeed, judges' behaviour implies that in an economic downturn, effective firing costs are lower: this would all other things equal encourage firms to fire and hence amplify the economic cycle, at least up to the point where most dismissed workers stay unemployed[21], in which case worse economic conditions would again favour workers. This line of reasoning implies however that firms are aware of judges' sensitivity to economic conditions, which may not be the case. Indeed, we have shown that the distribution of firms' settlement offers is unaffected by economic conditions, whereas one would expect that, all other things equal, firms lower their offers when economic conditions are worse, reflecting the lower probability of the workers' winning. However, only firms with personnel departments seem to be somewhat more likely to make no offer when economic conditions are degraded, which implies that only the most informed among firms may actually be aware of judges' decision patterns. Although most firms in our sample do not seem to be aware of the dependence of judges' decisions on economic conditions, if firms who do realize it are responsible for an important fraction of the dismissals, then judges' behaviour can have a macroeconomic effect. This study thus suggests that one should include indicators of EPL enforcement, such as workers' winning rate, in any study of the effect of EPL on macroeconomic outcomes.

An interesting avenue for future research would be to extend the analysis to countries such as the United States who, while not possessing any widespread dismissal legislation, operate similar institutions. Specifically, in the United States, one could investigate whether economic conditions influence committees and judges when deciding on appeals against unemployment benefits disqualification, or arbitrators when deciding about the regularity of dismissals in unionized firms. Given the

[21] A back of the envelope computation shows that one would need an unemployment rate of more than 20% to reverse judges' overall tendency to decide in favor of firms as the unemployment rate gets higher. However, such unemployment rates were never observed in the UK outside major depressions, so that overall judges' behavior, even in a high unemployment context, favors firms.

similarity in institutions, one would expect to find similar results to those found for the decisions of judges in labour courts in the United Kingdom and France.

To get a deeper understanding of the causes and consequences of the results reported in this paper, one should also examine to what extent firms and workers are aware of judges' decision rules and how such awareness affects their decisions. This, combined with a closer analysis of the reasons behind judges' sensitivity to economic conditions, would allow further examining whether judges' behaviour is efficient in maximizing social welfare and, armed with this knowledge, suggesting some suitable changes in the regulation. More generally, examining the influence of the socio-economic context in judges' decisions in other areas of law would likely permit to uncover interesting yet undiscovered patterns.

# Bibliography


ASHENFELTER O., BLOOM D. E. (1984), « Models of Arbitrators Behavior: Theory and Evidence », *The American Economic Review*, Vol. 74, No. 1 (Mar., 1984), 111-124.

AUTOR D. H., DONOHUE III J. J., SCHWAB S. J. (2002), « The costs of wrongful discharge laws », NBER Working Paper 9425, December 2002.

BEBCHUK A. B. (1984), « Litigation and Settlement under Imperfect Information », *The RAND Journal of Economics*, Vol. 15, No. 3 (Autumn, 1984), 404-415.

BENTOLILA S., BERTOLA G. (1990), « Firing Costs and Labour Demand : How Bad is Eurosclerosis ? », *The Review of Economic Studies,* Vol. 57, Issue 3 (Jul. 1990), 381-402.

BERTOLA G. (1992), "Labor Turnover Costs and Average Labor Demand," *Journal of Labor Economics*, 10:4 (October 1992), pp.389-411.

BERTOLA G. (2002), "A pure theory of Job Security and Labor Income Risk," *Review of Economic Studies,* forthcoming.

BERTOLA G., BOERI T., CAZES S. (1999), "Employment protection and labour market adjustment in OECD countries: Evolving institutions and variable enforcement", 1999.

BERTOLA G., CABALLERO R. (1994), « Cross-sectional Efficiency and Labour Hoarding in a Matching Model of Unemployment », *The Review of Economic Studies*, Vol. 61, Issue 3 (Jul., 1994), 435-456.

BOERI, T. (1999) « Enforcement of Employment Security Regulations, On-the-job Search and Unemployment Duration », *European Economic Review*, 43, 65-89.

BOERI T., GARIBALDI P., MACIS M., MAGGIONI M. (2002), « Adaptability of labour markets : a tentative definition and a synthetic indicator », Fondazione Rodolfo Debenedetti, Milan, Juin 2002.

BURGESS S. (1992), « Nonlinear Dynamics in a Structural Model of Employment », *Journal of Applied Econometrics*, Dec 1992.

BURGESS S., KNETTER M., MICHELACCI C. (2000), « Employment and Output Adjustment in the OECD: A Disaggregate Analysis of the Role of Job Security Provisions », *Economica*, Vol 67, 2000.

BURGESS S., PROPPER C., WILSON D. (2001), « Explaining the Growth in the Number of Applications to Indutrial Tribunals, 1972-1997 », Department of Trade and Industry and University of Bristol, 2001.

CABALLERO R., HAMMOUR M. (1996), « On the Timing and Efficiency of Creative Destruction », *Quarterly Journal of Economics* 1996.

COOTER, R., RUBINFELD D., *Journal of Economic Literature*, Vol. 27, No. 3. (Sep., 1989), pp. 1067-1097.

DAVIES P.L., FREEDLAND M. (1993), *Labour legislation and public policy: a contemporary history.* Oxford: Clarendon Press, 1993.

DAVIS S., HALTIWANGER J. (1992), « Gross Job Creation, Gross Job Destruction, and Employment Reallocation », *Quarterly Journal of Economics,* Volume 107, Issue 3, Aug. 1992.

DAUGHETY A. F. (1999), « Settlement ».

DICKENS L., HART M., JONES M., WEEKES B. (1984), « The British Experience under a Statute Prohibiting Unfair Dismissal », *Industrial and Labor Relations Review*, Volume 37, Issue 4 (Jul. 1984), 497-514.

DJANKOV S., LA PORTA R., LOPEZ-DE-SILANE F., SHLEIFER A., BOTERO J. (2003), « The Regulation of Labour», NBER Working Paper No. w9756, June 2003.

FARBER H. (1995), « Are Lifetime Jobs Disappearing ? Job Duration in the United States : 1973-1993 », NBER Working Paper No. 5014, 1995.

FARBER H. (1999), "Mobility and Stability: The Dynamics of Job Change in Labor Markets," The Handbook of Labor Economics, Orley Ashenfelter and David Card (eds.) Vol 3, Elsevier Science, 1999.

FARBER H. S., BAZERMAN M. H. (1986), « The General Basis of Arbitrator Behavior: An Empirical Analysis of Conventional and Final-Offer Arbitration », *Econometrica*, Vol. 54, Issue 6 (Nov. 1986), 1503-1528.

GUELL M. (2000), « Employment Protection and Unemployment in an Efficiency Wage Model », Center for Economic Performance, July 2000.

HAMERMESH D. (1995), « Labour Demand and the Source of Adjustement Costs », *The Economic Journal*, Volume 105, Issue 430, May 1995.

HAMERMESH D., PFANN G. (1996), « Adjustment Costs in Factor Demand », *Journal of Economic Literature*, Volume 34, Issue 3, Sep. 1996.

HECKMAN J.(1979), "Sample Selection Bias as a Specification Error", *Econometrica*, Vol. 47, No. 1. (Jan., 1979), pp. 153-162.

HUNT J. (2000), « Firing Costs, Employment Fluctuations and Average Employment : An Examination of Germany », *Economica* 2000, 67, 177-202.

ICHINO A., POLO M., RETTORE E. (2002), « Are Judges Biased by Labor Market Conditions ? », working paper, forthcoming in *European Economic Review*.

LAZEAR E. (1990), « Job Security Provisions and Employment », *Quarterly Journal of Economics*, Volume 105, Issue 3, Aug. 1990.

LJUNGQVIST L. (2002), « How Do Lay-off Costs Affect Employment », *The Economic Journal*, 112(October), 829-853.

MACIS M. (2001), « Il mercato del lavoro e la giurisprudenza in materia di licenziamenti (Italia, 1989-1998) », Diritto delle Relazioni Industriali N. 2, 2001.

MARINESCU I. (2003), *L'application du droit du travail est-elle biaisée par les conditions économiques?*, paper for the ADRES conference in Paris, january 2003.

MALCOMSON J. (1999), "Individual Employment Contracts," The Handbook of Labor Economics, Orley Ashenfelter and David Card (eds.) Vol 3, Elsevier Science, 1999.

MORTENSEN D., PISSARIDES C. (1994), « Job Creation and Job Destruction in the Theory of Unemployment », *The Review of Economic Studies*, Volume 61, Issue 3, Jul.1994.

OECD (1999), « EMPLOYMENT PROTECTION AND LABOUR MARKET PERFORMANCE », EMPLOYMENT OUTLOOK, 1999.

PFANN G., PALM F. (1993), « Asymmetric Adjustment Costs in Non-linear Labour Demand Models for the Netherlands and U.K. Manufacturing Sectors », *The Review of Economic Studies*, Volume 60, Issue 2, Apr. 1993.

REINGANUM J. F., WILDE L.L. (1986), « Settlement, Litigation, and the Allocation of Litigation Costs », *The RAND Journal of Economics*, Vol. 17, No. 4 (Winter, 1986), 557-566.

SARTORI A. (2003), "An Estimator for Some Binary-Outcome Selection Models Without Exclusion Restrictions", *Political Analysis* 2003 11: 111-138

SHAPIRO C., STIGLITZ J. (1984), *The American Economic Review*, Vol. 74, No. 3. (Jun., 1984), pp. 433-444.

TREMLETT N., BANERJI N. (1994), «The1992 survey of Industrial Tribunal Applications », dataset available at the UK Data Archive (http://www.data-archive.ac.uk).

VAN DE VEN W., VAN PRAAG B. (1981), "The demand for deductibles in private health insurance : A probit model with sample selection", *Journal of Econometrics*, Volume 17, Issue 2, November 1981, 229-252.

WALDFOGEL (1998), « Reconciling asymmetric information and divergent expectations theories of litigation », *Journal of Law and Economics*, Vol. XLI (Oct. 1998), 451-476

**Table 2 Descriptive statistics (no employment status)**

| *Variable* | *All applicants* | | | | | *Applicants proceeding to trial* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Obs* | *Mean* | *Std. Dev.* | *Min* | *Max* | *Obs* | *Mean* | *Std. Dev.* | *Min* | *Max* |
| Judges' decision | | | | | | | | | | |
| Worker wins | 471 | 0.43 | 0.50 | 0 | 1 | 471 | 0.43 | 0.50 | 0 | 1 |
| Economic conditions | | | | | | | | | | |
| Unemployment rate (%) | 1288 | 6.03 | 1.67 | 2.40 | 10.10 | 459 | 6.01 | 1.62 | 2.40 | 10.10 |
| Bankruptcy (deregistration) rate (%) | 1289 | 0.12 | 0.02 | 0.03 | 0.18 | 459 | 0.12 | 0.02 | 0.04 | 0.18 |
| CONTROLS: X | | | | | | | | | | |
| Case characteristics | | | | | | | | | | |
| Settlement offer/legal award | 1073 | 0.22 | 0.51 | 0 | 5.20 | 392 | 0.05 | 0.29 | 0 | 3.61 |
| Severe misconduct | 1311 | 0.15 | 0.36 | 0 | 1 | 471 | 0.18 | 0.38 | 0 | 1 |
| Economic dismissal | 1311 | 0.29 | 0.45 | 0 | 1 | 471 | 0.25 | 0.43 | 0 | 1 |
| Redundancy payment | 1311 | 0.09 | 0.29 | 0 | 1 | 471 | 0.09 | 0.29 | 0 | 1 |
| Internal formal procedure followed | 1311 | 0.29 | 0.46 | 0 | 1 | 471 | 0.34 | 0.47 | 0 | 1 |
| Firms' settlement offer (thousands of pounds) | 1311 | 0.75 | 1.98 | 0 | 30 | 471 | 0.13 | 0.71 | 0 | 10 |
| Worker characteristics | | | | | | | | | | |
| Manager or professional | 1311 | 0.21 | 0.40 | 0 | 1 | 471 | 0.24 | 0.42 | 0 | 1 |
| Weekly wage (hundreds of pounds) | 1173 | 2.02 | 1.25 | 0.15 | 15.38 | 423 | 2.11 | 1.29 | 0.30 | 12.50 |
| Tenure at dismissal (years) | 1278 | 7.00 | 6.45 | 0.08 | 41.00 | 460 | 7.42 | 6.50 | 0.08 | 35.00 |
| Age (tens of years) | 1227 | 4.02 | 1.19 | 1.70 | 7.10 | 445 | 4.11 | 1.15 | 1.80 | 6.40 |
| Female | 1311 | 0.32 | 0.47 | 0 | 1 | 471 | 0.28 | 0.45 | 0 | 1 |
| Firm characteristics | | | | | | | | | | |
| Size (hundreds of employees) | 1271 | 2.30 | 9.20 | 0.01 | 240 | 459 | 2.47 | 7.71 | 0.01 | 80 |
| Personnel department | 1311 | 0.18 | 0.39 | 0 | 1 | 471 | 0.21 | 0.41 | 0 | 1 |

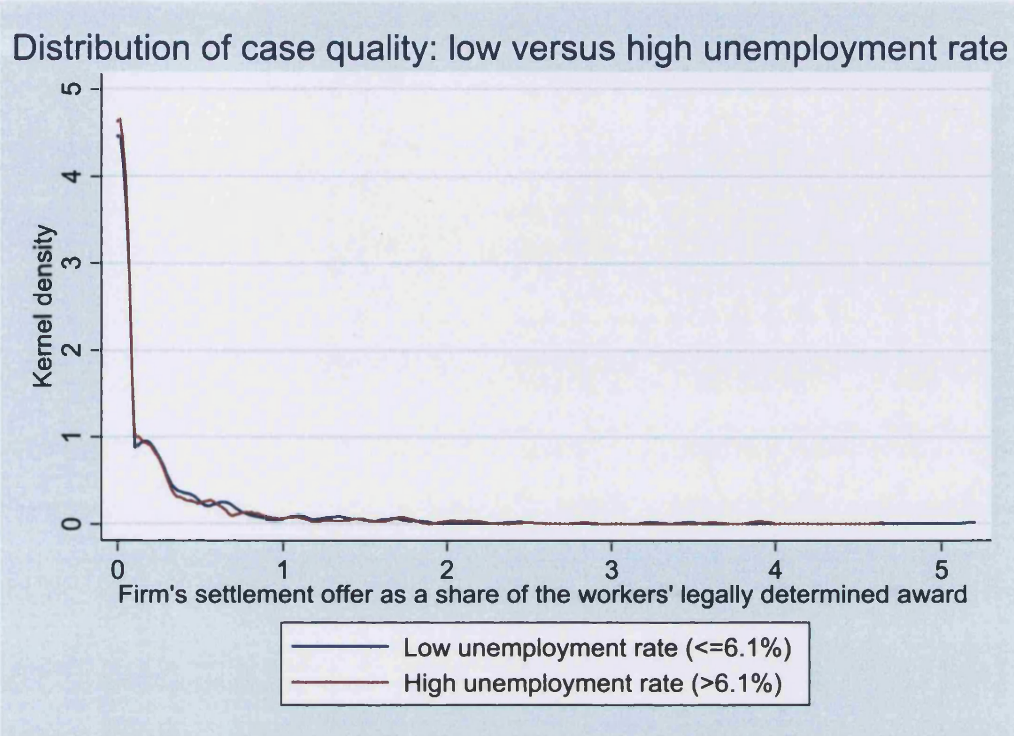
Source: 1992 survey of Employment Tribunal Applications in Great Britain, UK National Statistics, claimant count series and Small Business Service, VAT Deregistration.

**Table 3 Descriptive statistics (with employment status)**

| | All applicants | | | | | Applicants proceeding to trial | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *Variable* | *Obs* | *Mean* | *Std. Dev.* | *Min* | *Max* | *Obs* | *Mean* | *Std. Dev.* | *Min* | *Max* |
| **Judges' decision** | | | | | | | | | | |
| **Worker wins** | 135 | 0.36 | 0.48 | 0 | 1 | 135 | 0.36 | 0.48 | 0 | 1 |
| **Economic conditions** | | | | | | | | | | |
| **Unemployment rate (%)** | 358 | 6.05 | 1.78 | 2.40 | 10.10 | 131 | 5.94 | 1.74 | 2.40 | 10.10 |
| **Bankruptcy (deregistration) rate (%)** | 358 | 0.13 | 0.02 | 0.04 | 0.18 | 131 | 0.13 | 0.02 | 0.04 | 0.18 |
| CONTROLS: X | | | | | | | | | | |
| **Case characteristics** | | | | | | | | | | |
| **Settlement offer/legal award** | 304 | 0.19 | 0.40 | 0 | 3.90 | 111 | 0.02 | 0.11 | 0 | 0.83 |
| **Severe misconduct** | 365 | 0.16 | 0.37 | 0 | 1 | 135 | 0.16 | 0.37 | 0 | 1 |
| **Economic dismissal** | 365 | 0.33 | 0.47 | 0 | 1 | 135 | 0.26 | 0.44 | 0 | 1 |
| **Redundancy payment** | 365 | 0.01 | 0.12 | 0 | 1 | 135 | 0.02 | 0.15 | 0 | 1 |
| **Internal formal procedure followed** | 365 | 0.31 | 0.46 | 0 | 1 | 135 | 0.34 | 0.48 | 0 | 1 |
| **Firms' settlement offer (thousands of pounds)** | 365 | 0.72 | 1.71 | 0 | 15 | 135 | 0.08 | 0.45 | 0 | 4.50 |
| **Worker characteristics** | | | | | | | | | | |
| **Worker unemployed** | 365 | 0.23 | 0.42 | 0 | 1 | 135 | 0.26 | 0.44 | 0 | 1 |
| **Manager or professional** | 365 | 0.23 | 0.42 | 0 | 1 | 135 | 0.23 | 0.42 | 0 | 1 |
| **Weekly wage (hundreds of pounds)** | 328 | 2.05 | 1.14 | 0.24 | 10.58 | 122 | 2.12 | 1.26 | 0.35 | 10.58 |
| **Tenure at dismissal (years)** | 361 | 7.17 | 6.56 | 0.08 | 40 | 133 | 6.97 | 6.14 | 0.08 | 30.58 |
| **Age (tens of years)** | 346 | 3.94 | 1.16 | 1.70 | 7.10 | 126 | 3.98 | 1.14 | 1.80 | 6.20 |
| **Female** | 365 | 0.31 | 0.46 | 0 | 1 | 135 | 0.31 | 0.46 | 0 | 1 |
| **Firm characteristics** | | | | | | | | | | |
| **Size (hundreds of employees)** | 361 | 1.92 | 6.36 | 0.01 | 75 | 135 | 2.13 | 6.58 | 0.01 | 48 |
| **Personnel department** | 365 | 0.16 | 0.37 | 0 | 1 | 135 | 0.18 | 0.38 | 0 | 1 |

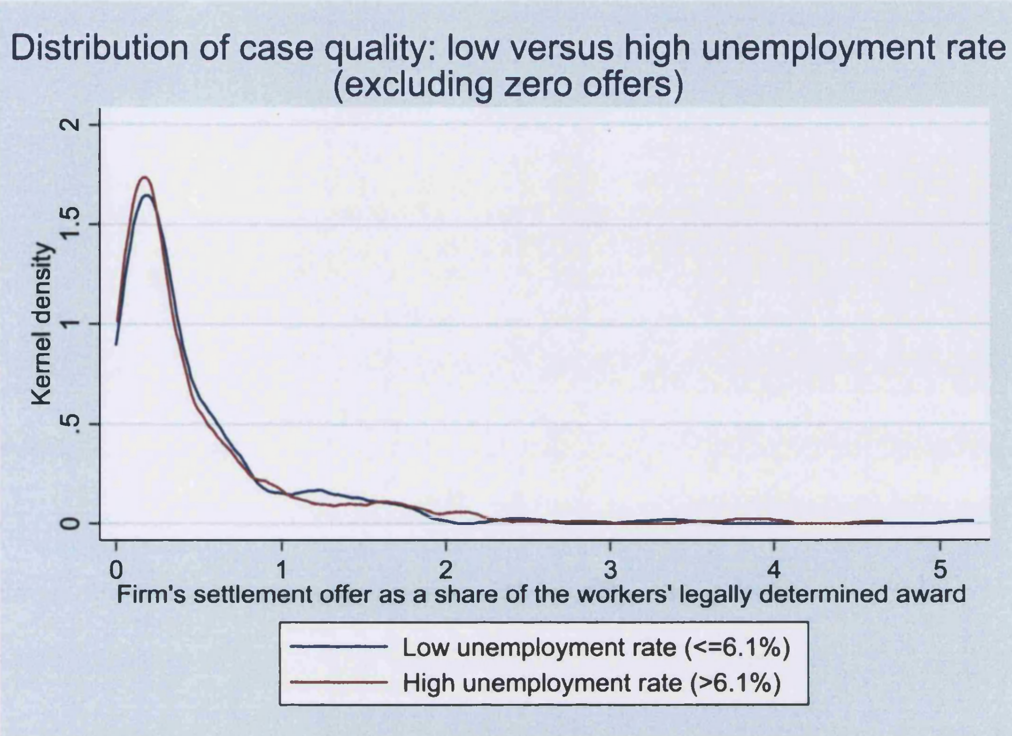
Source: 1992 survey of Employment Tribunal Applications in Great Britain, UK National Statistics, claimant count series and Small Business Service, VAT Deregistration.

**Figure 4: Distribution of case quality: low versus high unemployment rate**



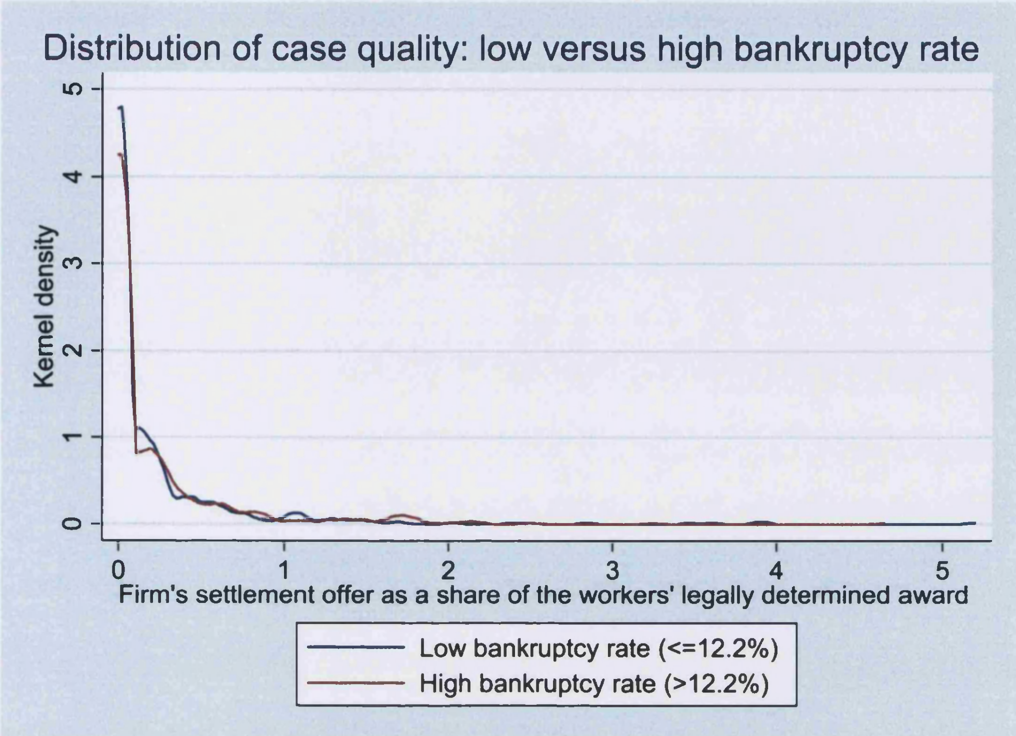

Source: 1992 survey of Employment Tribunal Applications in Great Britain, UK National Statistics, claimant count series.

**Figure 5: Distribution of case quality: low versus high unemployment rate (excluding zero offers)**



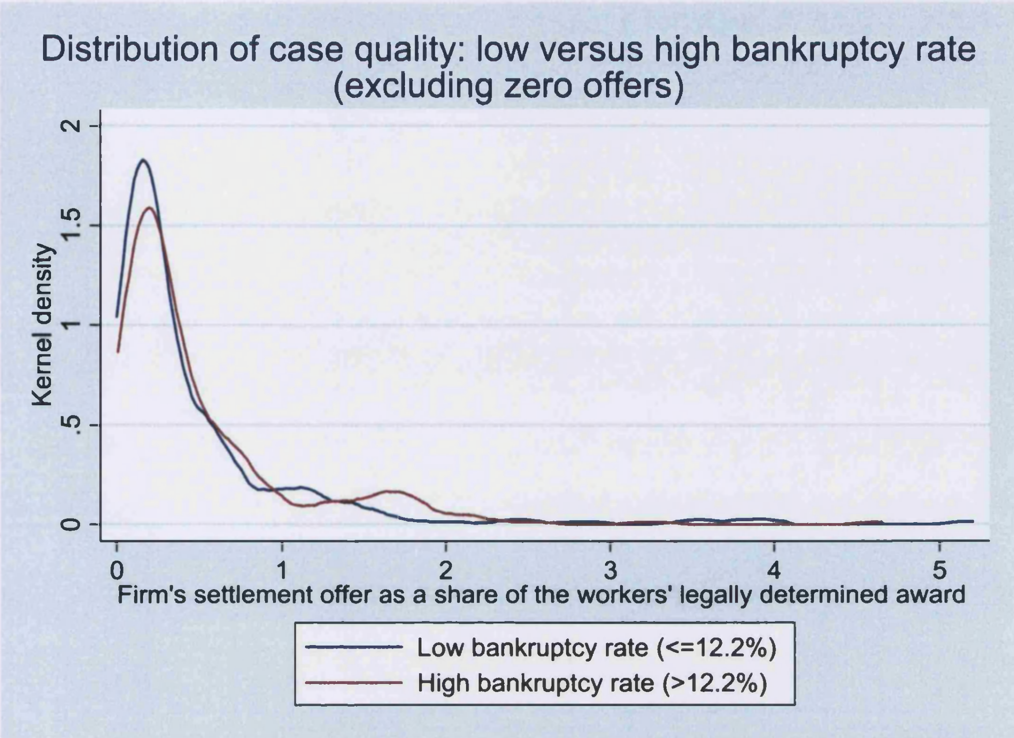

Source: 1992 survey of Employment Tribunal Applications in Great Britain, UK National Statistics, claimant count series.

**Figure 6: Distribution of case quality: low versus high bankruptcy rate**



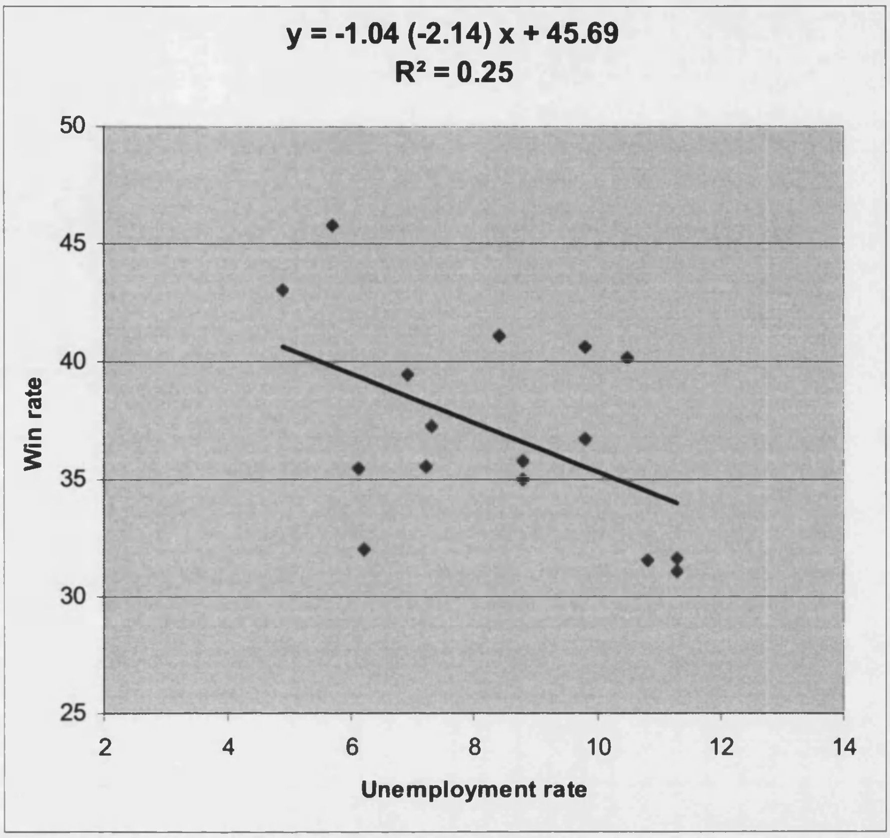

Source: 1992 survey of Employment Tribunal Applications in Great Britain, and Small Business Service, VAT Deregistration.

**Figure 7: Distribution of case quality: low versus high bankruptcy rate (excluding zero offers)**




Source: 1992 survey of Employment Tribunal Applications in Great Britain, and Small Business Service, VAT Deregistration.

**Figure 8: yearly win rate in unfair dismissal cases and unemployment rate (1985-2001)**




Source: Burgess et ali. (2001) and UK National Statistics

**Table 4: probit estimations for trial outcomes**

| | (1) P(win=1) | (2) P(win=1) | (3) P(win=1) | (4) P(win=1) | (5) P(win=1) | (6) P(win=1) | (7) P(win=1) | (8) P(win=1) |
|---|---|---|---|---|---|---|---|---|
| **Unemployment rate (%)** | -0.033 | -0.077 | -0.073 | -0.144 | -0.068 | -0.183 | -0.149 | -0.591 |
| | (0.013)*** | (0.044)* | (0.048) | (0.064)** | (0.031)** | (0.130) | (0.106) | (0.236)** |
| **Bankruptcy** | -0.016 | -0.012 | -0.027 | -0.036 | -0.001 | -0.007 | -0.106 | -0.156 |
| **(deregistration) rate (%)** | (0.008)* | (0.010) | (0.013)** | (0.016)** | (0.014) | (0.022) | (0.053)** | (0.050)*** |
| **Worker unemployed** | | | | | -0.638 | -0.683 | -0.825 | -0.796 |
| | | | | | (0.087)*** | (0.123)*** | (0.128)*** | (0.157)*** |
| **Unemployment rate*** | | | | | 0.132 | 0.157 | 0.168 | 0.212 |
| **worker unemployed** | | | | | (0.045)*** | (0.079)** | (0.077)** | (0.148)* |
| **Regional dummies** | No | Yes | Yes | Yes | No | Yes | Yes | Yes |
| **Industry dummies** | No | No | Yes | Yes | No | No | Yes | Yes |
| **Year dummy** | No | No | No | Yes | No | No | No | Yes |
| **Observations** | 387 | 387 | 387 | 387 | 112 | 108 | 93 | 93 |
| **Pseudo R squared** | 0.092 | 0.101 | 0.123 | 0.126 | 0.187 | 0.261 | 0.321 | 0.360 |

*** significant at 10%; ** significant at 5%; *** significant at 1%**

Notes: Robust standard errors clustered by region in parentheses. Marginal effects reported; the marginal effect of the interaction term is calculated using inteff. All regressions include controls for case characteristics (settlement offer/legal award, severe misconduct dummy, economic dismissal dummy, redundancy payment dummy, dummy for internal procedure having been followed, firm's settlement offer), worker characteristics (manager or professional dummy, weekly wage, tenure at dismissal, age, female dummy), and firm characteristics (size, dummy for personnel department). In columns 5 to 8, the sample is reduced because whether the worker is unemployed or not is only known for a subsample of cases (see text for more explanations).

Source: 1992 survey of Employment Tribunal Applications in Great Britain, UK National Statistics, claimant count series and Small Business Service, VAT Deregistration.

**Table 5: Heckman and Sartori estimations for trial selection and outcomes**

| | (1) Heckman | (2) Sartori | (3) Heckman | (4) Sartori |
|---|---|---|---|---|
| | Outcome equation: P(win=1/trial=1) | Outcome equation: P(win=1/trial=1) | Outcome equation: P(win=1/trial=1) | Outcome equation: P(win=1/trial=1) |
| **Unemployment rate (%)** | -0.034 | -0.042 | -0.075 | -0.049 |
| | (0.017)* | (0.020)** | (0.042)* | (0.026)** |
| **Bankruptcy** | -0.016 | -0.027 | -0.001 | -0.006 |
| **(deregistration) rate (%)** | (0.012) | (0.013)** | (0.026) | (0.018) |
| **Worker unemployed** | | | -0.624 | -0.697 |
| | | | (0.207)*** | (0.294)*** |
| **Unemployment rate*** | | | 0.179 (a) | 0.111(a) |
| **worker unemployed** | | | (0.098)* | (0.039)*** |
| | Selection equation: | Selection equation: | Selection equation: | Selection equation: |
| | P(trial=1) | P(trial=1) | P(trial=1) | P(trial=1) |
| **Unemployment rate (%)** | -0.007 | -0.006 | -0.007 | -0.007 |
| | (0.009) | (0.008) | (0.012) | (0.008) |
| **Bankruptcy** | -0.009 | -0.008 | -0.005 | -0.003 |
| **(deregistration) rate (%)** | (0.006) | (0.005)* | (0.008) | (0.006) |
| **Worker unemployed** | | | -0.011 | -0.028 |
| | | | (0.148) | (0.084) |
| **Unemployment rate*** | | | 0.0151 (b) | .0151 (b) |
| **worker unemployed** | | | (0.033) | (0.033) |
| **Rho** | 0.689 | | -0.689 | |
| | (0.323) | | (0.874) | |
| **Likelihood ratio test** | chi2(1)=0.04 | | chi2(1)=0.37 | |
| **of indep. eqns. (rho=0)** | Prob>chi2=0.844 | | Prob>chi2=0.542 | |
| **Observations** | 1063 | 1063 | 305 | 305 |

*** significant at 10%; ** significant at 5%; *** significant at 1%**

Notes: Robust standard errors clustered by region in parentheses. Marginal effects reported, except for (a). (b): marginal interaction effects calculated with inteff from a probit estimation of P(trial=1). All regressions include controls for case characteristics (settlement offer/legal award, severe misconduct dummy, economic dismissal dummy, redundancy payment dummy, dummy for internal procedure having been followed, firm's settlement offer), worker characteristics (manager or professional dummy, weekly wage, tenure at dismissal, age, female dummy), and firm characteristics (size, dummy for personnel department). In columns 3 and 4, the sample is reduced because whether the worker is unemployed or not is only known for a subsample of cases (see text for more explanations).
Source: 1992 survey of Employment Tribunal Applications in Great Britain, UK National Statistics, claimant count series and Small Business Service, VAT Deregistration.

# ANNEXES

**Table A-1: probit estimations for trial outcomes: full results**

| | (1) | | (2) | | (3) | | (4) | |
|---|---|---|---|---|---|---|---|---|
| | P(win=1) | | P(win=1) | | P(win=1) | | P(win=1) | |
| **Economic conditions** | | | | | | | | |
| Unemployment rate (%) | -0.033*** | (0.013) | -0.077* | (0.044) | -0.073 | (0.048) | -0.144** | (0.064) |
| Bankruptcy (deregistration) rate (%) | -0.016* | (0.008) | -0.012 | (0.010) | -0.027** | (0.013) | -0.036** | (0.016) |
| **Case characteristics** | | | | | | | | |
| Severe misconduct | -0.252*** | (0.058) | -0.250*** | (0.061) | -0.266*** | (0.066) | -0.260*** | (0.067) |
| Economic dismissal | -0.074 | (0.050) | -0.068 | (0.051) | -0.094* | (0.049) | -0.095** | (0.048) |
| Redundancy payment | 0.145 | (0.122) | 0.134 | (0.123) | 0.133 | (0.123) | 0.122 | (0.120) |
| Internal formal procedure followed | -0.091** | (0.044) | -0.097** | (0.045) | -0.104** | (0.048) | -0.102** | (0.050) |
| Firms' settlement offer (thousands of pounds) | 0.041* | (0.022) | 0.044* | (0.023) | 0.053** | (0.023) | 0.050** | (0.022) |
| **Worker characteristics** | | | | | | | | |
| Manager or professional | 0.101* | (0.057) | 0.098* | (0.058) | 0.086 | (0.066) | 0.093 | (0.069) |
| Weekly wage (hundreds of pounds) | -0.059** | (0.025) | -0.056** | (0.027) | -0.057** | (0.026) | -0.059** | (0.024) |
| Tenure at dismissal (years) | -0.005 | (0.005) | -0.005 | (0.006) | -0.004 | (0.006) | -0.004 | (0.006) |
| Age (tens of years) | -0.013 | (0.021) | -0.011 | (0.020) | -0.011 | (0.020) | -0.011 | (0.020) |
| Female | 0.063 | (0.053) | 0.051 | (0.060) | 0.036 | (0.062) | 0.036 | (0.063) |
| **Firm characteristics** | | | | | | | | |
| Size (hundreds of employees) | -0.001 | (0.004) | -0.001 | (0.005) | -0.001 | (0.005) | -0.001 | (0.005) |
| Personnel department | -0.115 | (0.073) | -0.117 | (0.082) | -0.133 | (0.085) | -0.130 | (0.086) |
| **Regional dummies (default=East Midlands)** | | | | | | | | |
| East | | | -0.063 | (0.050) | -0.075 | (0.056) | -0.162** | (0.079) |
| London | | | -0.010 | (0.037) | 0.008 | (0.057) | 0.027 | (0.061) |
| North East | | | 0.278 | (0.170) | 0.255 | (0.195) | 0.466*** | (0.146) |
| North West | | | 0.034 | (0.100) | 0.012 | (0.111) | 0.164 | (0.147) |
| South East | | | -0.151** | (0.060) | -0.158** | (0.069) | -0.267*** | (0.087) |
| South West | | | 0.046* | (0.024) | 0.078* | (0.042) | 0.027 | (0.055) |
| Scotland | | | 0.061 | (0.126) | 0.033 | (0.130) | 0.193 | (0.159) |
| West Midlands | | | 0.074 | (0.053) | 0.059 | (0.048) | 0.096** | (0.046) |
| Wales | | | -0.033 | (0.076) | -0.012 | (0.085) | 0.103 | (0.104) |
| Yorkshire and the Humber | | | 0.112 | (0.071) | 0.085 | (0.078) | 0.183** | (0.090) |
| **Industry dummies (default=Agriculture)** | | | | | | | | |
| Catering | | | | | 0.457*** | (0.104) | 0.498*** | (0.087) |
| Construction | | | | | 0.244* | (0.127) | 0.294** | (0.129) |
| Finance | | | | | 0.212*** | (0.068) | 0.250*** | (0.072) |
| Other services | | | | | 0.286** | (0.118) | 0.362*** | (0.129) |
| Production | | | | | 0.252** | (0.109) | 0.307*** | (0.111) |
| Retail | | | | | 0.081 | (0.123) | 0.153 | (0.125) |
| Transport | | | | | 0.105 | (0.125) | 0.184 | (0.124) |
| Wholesale | | | | | 0.276 | (0.184) | 0.346** | (0.173) |
| **Year dummy (default=1990)** | | | | | | | | |
| 1991 | | | | | | | 0.123* | (0.072) |
| Observations | 387 | | 387 | | 387 | | 387 | |
| Pseudo R squared | 0.092 | | 0.101 | | 0.123 | | 0.126 | |

* significant at 10%; ** significant at 5%; *** significant at 1%

Notes: Robust standard errors clustered by region in parentheses. Marginal effects reported.
Source: 1992 survey of Employment Tribunal Applications in Great Britain, UK National Statistics, claimant count series and Small Business Service, VAT Deregistration.

**Table A-2: : probit estimations for trial outcomes controlling for employment status: full results**

| | (1) | | (2) | | (3) | | (4) | |
|---|---|---|---|---|---|---|---|---|
| | P(win=1) | | P(win=1) | | P(win=1) | | P(win=1) | |
| **Economic conditions and employment status** | | | | | | | | |
| **Unemployment rate (%)** | -0.068 | (0.031)** | -0.183 | (0.130) | -0.149 | (0.106) | -0.591 | (0.236)** |
| **Bankruptcy (deregistration) rate (%)** | -0.001 | (0.014) | -0.007 | (0.022) | -0.106 | (0.053)** | -0.156 | (0.050)*** |
| **Worker unemployed** | -0.638 | (0.087)*** | -0.683 | (0.123)*** | -0.825 | (0.128)*** | -0.796 | (0.157)*** |
| **Unemployment rate*worker unemployed** | 0.132 | (0.045)*** | 0.157 | (0.079)** | 0.168 | (0.077)** | 0.212 | (0.148)* |
| **Case characteristics** | | | | | | | | |
| **Severe misconduct** | -0.263 | (0.066)*** | -0.303 | (0.081)*** | -0.463 | (0.064)*** | -0.419 | (0.089)*** |
| **Economic dismissal** | -0.045 | (0.116) | -0.097 | (0.141) | -0.087 | (0.203) | -0.052 | (0.202) |
| **Redundancy payment** | -0.133 | (0.188) | -0.232 | (0.135)* | 0.025 | (0.337) | 0.163 | (0.376) |
| **Internal formal procedure followed** | -0.123 | (0.088) | -0.104 | (0.120) | -0.142 | (0.140) | -0.136 | (0.151) |
| **Firms' settlement offer (thousands of pounds)** | -0.019 | (0.174) | 0.008 | (0.186) | 0.241 | (0.265) | 0.212 | (0.250) |
| **Worker characteristics** | | | | | | | | |
| **Manager or professional** | 0.026 | (0.124) | 0.027 | (0.133) | -0.098 | (0.172) | -0.141 | (0.164) |
| **Weekly wage (hundreds of pounds)** | -0.100 | (0.054)* | -0.159 | (0.080)** | -0.133 | (0.071)* | -0.143 | (0.060)** |
| **Tenure at dismissal (years)** | -0.025 | (0.012)** | -0.025 | (0.015)* | -0.041 | (0.017)** | -0.042 | (0.018)** |
| **Age (tens of years)** | 0.013 | (0.037) | -0.001 | (0.047) | -0.022 | (0.049) | -0.020 | (0.041) |
| **Female** | -0.021 | (0.094) | -0.103 | (0.138) | -0.189 | (0.144) | -0.184 | (0.135) |
| **Firm characteristics** | | | | | | | | |
| **Size (hundreds of employees)** | -0.184 | (0.135) | -0.006 | (0.008) | -0.004 | (0.011) | 0.001 | (0.014) |
| **Personnel department** | -0.180 | (0.122) | -0.081 | (0.182) | 0.086 | (0.255) | -0.002 | (0.285) |
| **Regional dummies (default=East Midlands)** | | | | | | | | |
| **East** | | | -0.134 | (0.156) | 0.202 | (0.286) | -0.322 | (0.162)** |
| **London** | | | 0.267 | (0.158)* | 0.345 | (0.172)** | 0.449 | (0.124)*** |
| **North East** | | | 0.736 | (0.169)*** | 0.623 | (0.230)*** | 0.856 | (0.077)*** |
| **North West** | | | 0.308 | (0.320) | 0.256 | (0.229) | 0.809 | (0.106)*** |
| **South East** | | | 0.094 | (0.345) | 0.176 | (0.358) | -0.393 | (0.107)*** |
| **South West** | | | 0.336 | (0.128)*** | 0.269 | (0.142)* | -0.007 | (0.196) |
| **Scotland** | | | 0.076 | (0.322) | | | | |
| **West Midlands** | | | 0.354 | (0.126)*** | -0.059 | (0.129) | 0.209 | (0.138) |
| **Yorkshire and the Humber** | | | 0.470 | (0.145)*** | 0.431 | (0.107)*** | 0.753 | (0.065)*** |
| **Industry dummies (default=Agriculture)** | | | | | | | | |
| **Construction** | | | | | -0.518 | (0.029)*** | -0.475 | (0.141)*** |
| **Finance** | | | | | -0.520 | (0.044)*** | -0.457 | (0.176)*** |
| **Other services** | | | | | -0.612 | (0.059)*** | -0.338 | (1.062) |
| **Production** | | | | | -0.945 | (0.032)*** | -0.444*** | (0.147) |
| **Retail** | | | | | -0.575 | (0.057)*** | -0.450 | (0.414) |
| **Wholesale** | | | | | -0.476 | (0.025)*** | -0.362 | (0.504) |
| **Year dummy (default=1990)** | | | | | | | | |
| **1991** | | | | | | | 0.651 | (0.254)** |
| **Observations** | 112 | | 108 | | 93 | | 93 | |
| **Pseudo R squared** | 0.187 | | 0.261 | | 0.321 | | 0.360 | |

*** significant at 10%; ** significant at 5%; *** significant at 1%**

Notes: Robust standard errors clustered by region in parentheses. Marginal effects reported; the marginal effect of the interaction term is calculated using inteff. The sample is reduced compared to table A-1 because whether the worker is unemployed or not is only known for a subsample of cases (see text for more explanations).

Source: 1992 survey of Employment Tribunal Applications in Great Britain, UK National Statistics, claimant count series and Small Business Service, VAT Deregistration.

**Table A-3: Heckman and Sarotri estimations for trial selection and outcomes**

| | (1) Heckman | | (2) Sartori | | (3) Heckman | | (4) Sartori | |
|---|---|---|---|---|---|---|---|---|
| | Outcome equation: P(win=1/trial=1) | | Outcome equation: P(win=1/trial=1) | | Outcome equation: P(win=1/trial=1) | | Outcome equation: P(win=1/trial=1) | |
| **Economic conditions and employment status** | | | | | | | | |
| **Unemployment rate (%)** | -0.034 | (0.017)* | -0.042 | (0.020)** | -0.075 | (0.042)* | -0.049 | (0.026)** |
| **Bankruptcy (deregistration) rate (%)** | -0.016 | (0.012) | -0.027 | (0.013)** | -0.001 | (0.026) | -0.006 | (0.018) |
| **Worker unemployed** | | | | | -0.624 | (0.207)*** | -0.697 | (0.294)*** |
| **Unemployment rate*worker unemployed** | | | | | 0.179 (a) | (0.098)* | 0.111(a) | (0.039)*** |
| **Case characteristics** | | | | | | | | |
| **Severe misconduct** | -0.262 | (0.067)*** | -0.247 | (0.099)*** | -0.271 | (0.131)** | -0.172 | (0.119)* |
| **Economic dismissal** | -0.073 | (0.067) | -0.105 | (0.077)* | -0.047 | (0.136) | -0.050 | (0.098) |
| **Redundancy payment** | 0.147 | (0.099) | 0.118 | (0.109) | -0.136 | (0.273) | -0.020 | (0.281) |
| **Internal formal procedure followed** | -0.097 | (0.062) | -0.088 | (0.075) | -0.144 | (0.127) | -0.114 | (0.088)* |
| **Firms' settlement offer (thousands of pounds)** | 0.022 | (0.055) | -0.288 | (0.068)*** | -0.059 | (0.235) | -0.529 | (0.131)*** |
| **Worker characteristics** | | | | | | | | |
| **Manager or professional** | 0.094 | (0.074) | 0.200 | (0.088)** | 0.029 | (0.156) | 0.059 | (0.110) |
| **Weekly wage (hundreds of pounds)** | -0.063 | (0.029)** | -0.044 | (0.033)* | -0.101 | (0.075) | -0.050 | (0.049) |
| **Tenure at dismissal (years)** | -0.006 | (0.005) | -0.005 | (0.006) | -0.029 | (0.017)* | -0.020 | (0.009)** |
| **Age (tens of years)** | -0.014 | (0.026) | 0.027 | (0.029) | 0.024 | (0.061) | 0.051 | (0.037)* |
| **Female** | 0.064 | (0.067) | 0.032 | (0.075) | -0.018 | (0.130) | 0.043 | (0.091) |
| **Firm characteristics** | | | | | | | | |
| **Size (hundreds of employees)** | -0.001 | (0.005) | 0.001 | (0.006) | 0.000 | (0.009) | 0.001 | (0.007) |
| **Personnel department** | -0.12 | (0.078) | -0.093 | (0.100) | -0.199 | (0.158) | -0.163 | (0.145) |
| **Constant** | 0.705 | (0.717) | 0.167 | (0.659) | 2.427 | (1.078)** | 0.441 | (0.838) |

**Table A-3: Heckman and Sarotri estimations for trial selection and outcomes (continued)**

| | (1) Heckman | | (2) Sartori | | (3) Heckman | | (4) Sartori | |
|---|---|---|---|---|---|---|---|---|
| | Selection equation: P(trial=1) | | Selection equation: P(trial=1) | | Selection equation: P(trial=1) | | Selection equation: P(trial=1) | |
| Economic conditions and employment status | | | | | | | | |
| Unemployment rate (%) | -0.007 | (0.009) | -0.006 | (0.008) | -0.007 | (0.012) | -0.007 | (0.008) |
| Bankruptcy (deregistration) rate (%) | -0.009 | (0.006) | -0.008 | (0.005)* | -0.005 | (0.008) | -0.003 | (0.006) |
| Worker unemployed | | | | | -0.011 | (0.148) | -0.028 | (0.084) |
| Unemployment rate*worker unemployed | | | | | .0151 (b) | (0.033) | .0151 (b) | (0.033) |
| Case characteristics | | | | | | | | |
| Severe misconduct | 0.032 | (0.043) | 0.026 | (0.036) | 0.039 | (0.056) | 0.021 | (0.034) |
| Economic dismissal | -0.026 | (0.035) | -0.022 | (0.031) | -0.013 | (0.044) | -0.009 | (0.029) |
| Redundancy payment | -0.007 | (0.053) | -0.007 | (0.048) | 0.073 | (0.174) | 0.035 | (0.090) |
| Internal formal procedure followed | 0.013 | (0.033) | 0.011 | (0.028) | -0.028 | (0.038) | -0.016 | (0.026) |
| Firms' settlement offer (thousands of pounds) | -0.19 | (0.017)*** | -0.165 | (0.016)*** | -0.350 | (0.033)*** | -0.221 | (0.037) |
| Worker characteristics | | | | | | | | |
| Manager or professional | 0.096 | (0.044)** | 0.082 | (0.035)** | 0.018 | (0.057) | 0.003 | (0.039) |
| Weekly wage (hundreds of pounds) | 0.013 | (0.014) | 0.009 | (0.012) | 0.024 | (0.021) | 0.016 | (0.013) |
| Tenure at dismissal (years) | 0.002 | (0.003) | 0.002 | (0.002) | -0.001 | (0.003) | -0.001 | (0.002) |
| Age (tens of years) | 0.029 | (0.014)** | 0.025 | (0.012)** | 0.033 | (0.018)* | 0.018 | (0.012) |
| Female | -0.013 | (0.034) | -0.01 | (0.030) | 0.045 | (0.049) | 0.026 | (0.029) |
| Firm characteristics | | | | | | | | |
| Size (hundreds of employees) | 0.002 | (0.002) | 0.002 | (0.002) | 0.002 | (0.003) | 0.001 | (0.002) |
| Personnel department | 0.023 | (0.043) | 0.019 | (0.036) | -0.014 | (0.049) | -0.007 | (0.034) |
| Constant | -0.158 | (0.329) | -0.156 | (0.636) | -0.282 | (0.683) | -0.111 | (0.681) |
| Rho | 0.6894 | (0.323) | | | -0.689 | (0.874) | | |
| Likelihood ratio test of indep. eqns. (rho=0) | chi2(1)=0.04 | | | | chi2(1)=0.37 | | | |
| | Prob>chi2=0.844 | | | | Prob>chi2=0.542 | | | |
| Observations | 1063 | | 1063 | | 305 | | 305 | |

* significant at 10%; ** significant at 5%; *** significant at 1%

Notes: Robust standard errors clustered by region in parentheses. Marginal effects reported, except for (a). (b): marginal interaction effects calculated with inteff from a probit estimation of P(trial=1). In columns 3 and 4, the sample is reduced because whether the worker is unemployed or not is only known for a subsample of cases (see text for more explanations).

Source: 1992 survey of Employment Tribunal Applications in Great Britain, UK National Statistics, claimant count series and Small Business Service, VAT Deregistration.

# Chapter 4

# Cyclical Budgetary Policy and Economic Growth: What Do We Learn from OECD Panel Data?*

Philippe Aghion[†] and Ioana Marinescu

**Abstract:** This paper uses yearly panel data on OECD countries to analyze the relationship between growth and the cyclicality of the budget deficit. We develop new yearly estimates of the countercyclicality of the budget deficit, and show that the budget deficit has become increasingly countercyclical in most OECD countries over the past twenty years. However, EMU countries did not become more countercyclical. Using panel specifications with country and year fixed effects, we show that: (i) an increase in financial development, a decrease in openness to trade, and the adoption of an inflation targeting regime move countries toward a more countercyclical budget deficit; (ii) a more countercyclical budget deficit has a positive and significant effect on economic growth, and this effect is larger when financial development is lower.

*This work owes a lot to Robert Barro who contributed abundant advice, and to the very helpful comments and editorial suggestions of Daron Acemoglu, Olivier Blanchard, Ken Rogoff, and Michael Woodford. We also thank Ricardo Caballero and Anil Kashyap for their useful discussions. At earlier stages this project benefited from fruitful conversations with Philippe Bacchetta, Tim Besley, Laurence Bloch, Elie Cohen, Philippe Moutot, Jean Pisani-Ferry, Romain Ranciere, and of our colleagues in the Institutions, Organizations and Growth group at the Canadian Institute for Advanced Research. We are very grateful to Ann Helfman, Julian Kolev and Anne-Laure Piganeau for outstanding research assistance. Finally, we thank Konrad Kording for his collaboration on the first stage analysis section and more specifically for helping us implement the MCMC methodology.

[†]Harvard University and NBER

# 1 Introduction

A common view among macroeconomists, is that there is a decoupling between macroeconomic policy (budget deficit, taxation, money supply) which should primarily affect price and income stability[1], and long-run economic growth which, if anything, should depend only upon structural characteristics of the economy (property right enforcement, market structure, market mobility and so forth). That macroeconomic policy should not be a key determinant of growth, is further hinted at by recent contributions such as Acemoglu et al (2004) and Easterly (2005), which argue that the correlation between macroeconomic volatility and growth (Acemoglu et al) or those between growth and macroeconomic variables (Easterly), become insignificant once one controls for institutions.

The question of whether macroeconomic policy does or does not affect (productivity) growth is not purely academic. In particular, it underlies the recent debate on the European Stability and Growth Pact as well as the criticisms against the European Central Bank for allegedly pursuing price stability at the expense of employment and growth.

In this paper we question that view by arguing that the cyclicality of the budget deficit is significant in explaining GDP growth, with a more countercyclical budgetary policy being more growth-enhancing the lower the country's level of financial development. We also identify economic factors that tend to be associated with more countercyclical policies. These results hold in a sample of OECD countries with comparable institutional environments.

[1]For example Lucas (1987) analyzes the welfare costs of income volatility in an economy with complete markets for individual insurance, taking the growth rate as given. Atkeson and Phelan (1994) analyze the welfare gains from countercyclical policy in an economy with incomplete insurance markets but no growth. Both find very small effects of volatility (or of countercyclical policies aimed at reducing it) on welfare.

The idea that cyclical macroeconomic policy might affect productivity growth, is suggested by previous work by Aghion, Angeletos, Banerjee and Manova (2006), henceforth AABM. The argument in AABM is that credit constrained firms have a borrowing capacity which is typically conditioned by current earnings (the factor of proportionality between earning and debt capacity is called credit multiplier, with a higher multiplier reflecting a higher degree of financial development in the economy). In a recession, current earnings are reduced, and so is firms' ability to borrow in order to maintain growth-enhancing investments (e.g in skills, structural capital, or R&D). To the extent that higher macroeconomic volatility translates into deeper recessions, it should affect firms' incentives to engage in such investments. This prediction finds empirical support, first in cross-country panel regressions by AABM who show on the basis of cross-country panel regressions that structural investments are more procyclical the lower the country's level of financial development; and second, in firm-level evidence by Berman et al (2007). Using French firm-level panel data on R&D investments and on credit constraints, Berman et al. show that: (i) the share of R&D investment over total investment is countercyclical without credit constraints; (ii) this share turns more procyclical when firms are credit constrained; (iii) this effect is only observed during down-cycle phases - i.e. in presence of credit constraints, R&D investment share plummets during recessions but doesn't increase proportionally during up-cycle periods[2].

These findings in turn suggest that countercyclical macroeconomic policies, with higher government investment or lower nominal interest rates

[2] As pointed out by several authors, some of these results may be biased because of an endogeneity problem which may come from the the potential simultaneous determination of sales and investment. BEAAC check the robustness of their results by instrumenting the variation in sales by an exchange rate exposure variable, which depends on exchange rate variations and firms' export status. This variable is strongly correlated with sales variation without being affected by investment decisions. Their results are robust to this instrumentation.

during recessions, may foster productivity growth by reducing the magnitude of the output loss induced by market failures (in particular by credit market imperfections) in a recession, which in turn should allow credit-constrained firms to preserve their growth-enhancing investments over the business cycle. For example, the government may decide to stimulate the demand for private firms' products by increasing spending. This could further increase firm's liquidity holdings and thus make it easier for them to face idiosyncratic liquidity shocks without having to sacrifice R&D or other types of longer-term growth-enhancing investments. On the other hand, in a recession, more workers face unemployment, so that their earnings are reduced. Government spending could help them overcome credit constraints either directly (social programs, etc.) or indirectly by fostering labor demand and therefore employment; this relaxation of credit constraints in turn would allow workers to make growth-enhancing investments in human capital, re-location, etc. The tighter the credit constraints faced by firms and workers, the more growth-enhancing such countercyclical policies should be.[3]

Our contribution in this paper is three-fold. It is first to compute and analyze the cyclicality of the budget deficit on a panel of OECD countries, that is, how the budget deficit responds to fluctuations in the output gap over time. Second, it is to investigate some potential determinants of the countercyclicality of the budget deficit. Third, it is to use these yearly panel data to assess the relationship between growth and the countercyclicality of budgetary policies at various levels of financial development. Our main findings can be summarized as follows: (i) the budget deficit has become

[3]That government intervention might increase aggregate efficiency in an economy subject to credit constraints and aggregate shocks, has already been pointed out by Holmstrom and Tirole (1998). Our analysis in this section can be seen as a first attempt to explore potential empirical implications of this idea for the relationship between growth and public spending over the cycle.

increasingly countercyclical in most OECD countries over the past twenty years, but this trend has been significantly less pronounced in the EMU; (ii) within countries, a more countercyclical budgetary policy is positively associated with a higher level of financial development, a lower level of openness, and the adoption of an inflation targeting regime; (iii) a more countercyclical budgetary policy has a greater positive impact on growth when financial development is lower. While we argue that our results likely reflect the causality from budgetary policy to growth, at the very least they document statistical relationships between macroeconomic variables that are consistent with the theory and microevidence on volatility, credit constraints and growth-enhancing investments.

While we do not know of any previous attempt at analyzing the growth effects of countercyclical budgetary policies, analyses of the determinants of the cyclicality of budgetary policies already exist in the literature. For example, Alesina and Tabellini (2005) argue that more corrupt democracies will tend to run a more procyclical fiscal policy. The idea is that, in good times, voters demand that the government cut taxes or provide more public services instead of reducing debt, because they cannot observe the debt reduction and can suspect the government of appropriating the rents associated with good economic conditions. In equilibrium, this leads to a more procyclical policy as the moral hazard problem worsens, in the sense that governments are more likely to divert public resources in booms. They also show that this mechanism tends to be more powerful in explaining the variation observed in the data than borrowing constraints alone. While Alesina and Tabellini (2005) are using a large sample of countries and explore cross-sectional variations, in this study we use panel analysis on OECD countries. This makes the use of corruption indices impractical for two reasons. First, there is almost no cross-sectional variation in corruption indices within the OECD. Second, there is even less variation of these indices across time for individual countries.

In a similar vein, Calderon et al. (2004) show that emerging market economies with better institutions are more able to conduct a countercyclical fiscal policy[4]. Their empirical analysis is based on the International Country Risk Guide. Although the variation in this indicator is limited across OECD countries and time, it presents somewhat more variation than corruption indexes[5].

Other papers such as Gali and Perotti (2003) and Lane (2003) focus, as we do, on OECD countries. Gali and Perotti investigate whether fiscal policy in the European Monetary Union (EMU) has become more procyclical after the Maastricht treaty. They find no evidence for such a development. They do find however that while there is a trend in the OECD towards a more countercyclical fiscal policy over time, the EMU is lagging behind that trend. Lane (2003) is probably the paper that comes closer to the analysis developed in the third section of our paper. Lane examines the cyclical behavior of fiscal policy within the OECD. He then uses trade openness, output volatility, output per capita, the size of the public sector and an index for political power dispersion to examine cross-country differences in cyclicality. The reason why power dispersion may play a role is taken from Lane and Tornell (1998): when multiple political groups compete for public spending, the latter may become more procyclical. No group wants

[4]There is also the paper by Talvi and Vegh (2000), where it is argued that high output volatility is most likely to generate a procyclical government spending. The idea is that running a budget surplus generates political pressures to spend more: the government therefore minimizes that surplus and becomes pro-cyclical. This movement is then accentuated by a volatile output, and therefore a volatile tax base.

[5]We have also used these indicators in our analysis. However, they typically have no significant effect on GDP growth over time in our sample. Moreover, as they are less widely available than our main variables of interest, their use considerably restricts the available sample, leading to less precise estimates. We have therefore decided not to use these indicators in the results reported here.

to let any substantial fiscal surplus subsist because they are afraid that this will not lead to debt repayment, but rather to other groups appropriating that surplus. Lane finds in particular evidence that GDP growth volatility, trade openness and political divisions lead to a more procyclical spending pattern, even though the effect of political divisions is not present for all categories of spending. We contribute to this literature by using yearly panel data to analyze the cyclicality of budgetary policy and its determinants within OECD countries, and we show that the degree of financial development is an important element to explain within country variations in such policies, while future or present EMU membership explains cross-country variations. Moreover, we show that inflation targeting is associated with a more countercyclical budgetary deficit.

Most closely related to our second stage analysis of the effect of countercyclical budgetary policy on growth, are Aghion-Angeletos-Banerjee-Manova (2005), henceforth AABM, and Aghion-Bacchetta-Ranciere-Rogoff (2006), henceforth ABRR. AABM develop a model to explain why macroeconomic volatility is more negatively correlated with productivity growth, the lower financial development, and they test this prediction using cross-country panel data. ABRR move from a closed real to an open monetary economy and show that a fixed nominal exchange rate regime or lower real exchange rate volatility are more positively associated with productivity growth, the lower financial development and the lower the ratio of real shocks to financial shocks.

The remaining part of the paper is organized as follows. Section 2 discusses the estimation of the countercyclicality of the budget deficit for each OECD country and each year covered by our panel data set. Section 3 uncovers some main determinants of the countercyclicality of the budget deficit. Section 4 regresses GDP per capita growth on financial development, the countercyclicality coefficients computed in section2, the interaction between the two, and a set of controls. Finally, Section 5 concludes.

# 2 The countercyclicality of the budget deficit in the cross-country panel

In this section we compute time varying measures of the cyclicality of budgetary policy in our cross-country panel, and compare the extent to which budgetary policy became more countercyclical over time in some countries than in others. A main finding is that budgetary policy in the US and the UK have become significantly more countercyclical over the past twenty years, whereas it has not in the EMU area.

## 2.1 Data

Panel data on GDP, the GDP gap (ygap), the GDP deflator, and government gross debt (ggfl) are taken from the OECD Economic Outlook annual series[6]. Our measure of budgetary policy is the first difference of debt divided by the GDP, which is the same as the budget deficit over GDP. Note that debt and other government data refer to general government. Financial development is measured by the ratio of private credit to GDP, and annual cross-country data for this measure of financial development can be drawn from the Levine database[7]. In this latter measure, private credit is all credit to private agents, and therefore includes credit to households. The "'average years of education in the population over 25 years old"' se-

[6]Codes in parenthesis indicate the names of variables in the dataset. Full documentation available at www.oecd.org. Data can be downloaded from sourceoecd.org for subscribers to that service.

[7]Data downloadable from Ross Levine's homepage.

ries is directly borrowed from the Barro-Lee dataset; this measure is only available every five years and has been linearly interpolated to obtain a yearly series. The openness variable is defined as exports and imports over GDP and data on it come from the Penn World Tables 6.1. The population growth, government share of GDP and investment share of GDP also come from the Penn World Tables 6.1. The inflation targeting dummy is defined using the dates when countries adopted inflation targeting, as summarized in Vega and Winkelried (2005). All nominal variables are deflated using the GDP deflator. Summary statistics can be found in Table 1. The sample is an unbalanced panel including the following countries: Australia, Austria, Belgium, Canada, Denmark, Spain, Finland, France, United Kingdom, Germany[8], Iceland, Italy, Japan, Netherlands, Norway, New Zealand, Portugal, Sweden, USA.

## 2.2 Public deficit and growth: the empirical challenge

We are interested in evaluating the impact of the cyclicality of the budget deficit on the growth of GDP per capita, and how this effect may depend on the degree of financial development. Our expectation is that a more countercyclical budget deficit is more likely to enhance growth when financial development is lower. Empirically, we wish to identify this effect from time variation of budgetary policy within countries. Figure 1 illustrates this idea for a hypothetical case: we distinguish between the situation where, in the base period $t-1$, financial development is low (upper panels), and the sit-

[8]All level variables are adjusted for the German reunification. The adjustment involves regressing each variable of interest on time and a constant in the ten years before 1991 (data based on West Germany only). We then use the estimated coefficients to predict the values for 1991 to 2000. We take the average ratio between actual and predicted values in the years 1991 to 2000. We use this ratio to proportionally adjust values before 1991.

uation where financial development is high (lower panels). We start with a baseline depicted in the left-hand side panels of Figure 1: the budget deficit is thus initially assumed to be pro-cyclical. The right-hand side panels of Figure 1 illustrate the growth response in period 2 after an increase in the countercyclicality of the budget deficit in period 1, such that the budget deficit becomes strongly countercyclical. If financial development is low, then trend growth in period 2 increases substantially (upper left panel in Figure 1). If, on the other hand, financial development is high , then trend growth increases by a smaller amount[9] (lower left panel of Figure 1).

*FIGURE* 1 *HERE*

Looking at Figure 1, the most obvious method one can think of to compute cyclicality is to regress the public deficit on the GDP growth using ordinary least squares (OLS) on the observations in period $t$. In practice, it seems more reasonable to regress the public deficit on the GDP gap (defined as $(GDP-GDP^*)/GDP^*$, where $GDP^*$ is the trend GDP) rather than the GDP growth. Indeed, the GDP gap is very much like a detrended measure of the GDP growth, and a forward-looking government's budgetary policy should respond to shortfalls from trend rather than to GDP growth per se (for a theory of why fiscal policy should depend on the GDP gap, see Barro(1979)).

This type of regression based approach to measure the cyclicality of fiscal policies is now common in the literature and can be found for example in Lane (2003) and Alesina and Tabellini (2005). However, the methods used in these papers give rise to only one (or a few) observation of cyclicality per country. Since we want to investigate the impact of time variation in

[9]The effect of a decrease in the countercyclicality of public deficit could become negative at high enough levels of financial development, if the government's deficit crowds out more efficient private borrowing and spending.

cyclicality, we need to compute for each country time-varying measures for the countercyclicality of budget deficit. Specifically, as we wish to use a yearly panel of countries, we need a measure of countercyclicality that varies yearly. This means that period $t-1$ and period $t$ in Figure 1 are reduced to one single year each! A regression is not defined for a single observation, so we must use observations from a few years in order to compute countercyclicality. The next subsection discusses what methods can be used to compute countercyclicality.

## 2.3 Econometric methods to compute countercyclicality

Generally, one would like estimate the following equation for each country $i$:

$$\frac{b_{it} - b_{i,t-1}}{y_{it}} = -a_{1it}y_{gap,it} + a_{2it} + \varepsilon_{it}, \text{where } \varepsilon_{it} \sim N(0, \sigma_\varepsilon^2), \qquad (1)$$

where $a_{1it}$ measures the countercyclicality of budgetary policy. Note that there is a minus sign in front of $y_{gap,it}$: when the economy is in a recession and the GDP gap is negative, the opposite of the GDP gap is positive, and so a positive $a_{1it}$ means that the budget deficit increases when the economy is in a recession, i.e. the budget deficit is countercyclical. Both $a_{1it}$ and the constant $a_{2it}$ [10] are both time-varying, which is why we write $a_{jit}$ to denote the coefficient on the variable $j$ in country $i$ at year $t$.

The variables in equation 1 are defined as follows:

- $b_{it}$ : gross government debt in country $i$ at year $t$

[10] The constant $a_{2it}$ can be interpreted as a measure of structural budgetary deficit: indeed, by construction it corresponds to the part of budget deficit that does not depend upon the business cycle.

- $y_{it}$ : the GDP in country $i$ and year $t$, in value

- $y_{gap,it}$ : the GDP gap in country $i$ and year $t$. It is computed as $(y_{it} - y_{it}^*)/y_{it}^*$, where $y_{it}^*$ is the prediction of $y_{it}$ using the Hodrick-Prescott filter. A lambda parameter of 25 was chosen, following OECD(1995). Note that the GDP gap computed by the OECD using a production function approach is also smoothed by a Hodrick-Prescott technique, so that in practice the difference between the OECD measure of the GDP gap and the measure used here is very limited: the correlation between the two variables is 77%. Our measure of the GDP gap is as expected positively correlated with the GDP per capita growth: the correlation is however not so strong at 36%.

Note that $b_{it} - b_{i,t-1}$ is exactly equal to the opposite of the budget balance, so that our left-hand side variable is equal to the budget deficit as a share of GDP, which we will simply refer to as "budget deficit". We now examine how the coefficients $a_{jit}$ can be estimated econometrically.

One way to implement this is to compute finite (for example 10-years) rolling window ordinary least squares estimates. The ten-year rolling window OLS method simply amounts to estimating the countercyclicality of the budget deficit $\frac{(b_{it}-b_{i,t-1})}{y_{it}}$ at year $t$ in country $i$ by running the following regression for each country $i$, and all possible years $\tau$:

$$\frac{b_{it} - b_{i,t-1}}{y_{it}} = -a_{1it}y_{gap,i\tau} + a_{2it} + \varepsilon_{i\tau}, \text{ for } \tau \in (t-5, t+4). \tag{2}$$

that is, one uses a ten year centered rolling window to estimate the countercyclicality of budget deficit at any date $t$. This method suffers however from serious shortcomings. First, by definition, we lose the first five years and the last four years of data for each country. Second, because the method involves estimating a coefficient by discarding at each time period one old observation and taking into account a new one, the coefficient can vary substantially when the new observation is very different from the one it replaces. This implies that the series may be jagged and affected by noise and transitory changes; moreover, a sudden jump in the series would not be coming from changes in the immediate neighborhood of date $t$, but from changes 5 years before and 4 years after.

To deal with the shortcomings of the 10-years rolling window method, one can use smoothing such that all observations are used for each year, but those observations closest to the reference year are given greater weight. The "local Gaussian-weighted ordinary least squares" method is one way of achieving this. It consists in computing the $a_{jit}$ coefficients by using all the observations available for each country $i$ and then performing one regression for each date $t$, where the observations are weighted by a Gaussian centered at date $t$ [11]

$$\begin{aligned} \frac{b_{it} - b_{i,t-1}}{y_{it}} &= -a_{1it} y_{gap,i\tau} + a_{2it} + \varepsilon_{i\tau}, \qquad (3) \\ \text{where } \varepsilon_{i\tau} &\sim N(0, \sigma^2 / w_t(\tau)) \text{ and } w_t(\tau) = \frac{1}{\sigma\sqrt{2\Pi}} \exp\left(-\frac{(\tau - t)^2}{2\sigma^2}\right). \end{aligned}$$

While the local Gaussian-weighted OLS method is less noisy than the 10-years rolling window method, it suffers from a similar shortcoming when it comes to testing the idea illustrated in Figure 1. Indeed, these two methods use observations from both the past and the future (previous years and future years) to calculate yearly countercyclicality. Ultimately, we want to look at the impact of year $t-1$ changes in countercyclicality

[11] In practice, we chose a value of 5 for $\sigma$. While this choice is somewhat arbitrary, changing this smoothing value slightly does not qualitatively affect the results.

on year $t$ growth, but if countercyclicality is computed using some future observations, then in practice we are examining the impact of both past and (some) future countercyclicality on growth. Thus, it is hard to be certain that year $t-1$ countercyclicality causes year $t$ growth, and reverse causality becomes a problem. One way to address this issue is to use longer lags of countercyclicality ($t-2$ or $t-3$ or $t-4$, etc.), but this requires us to assume that the effects of countercyclicality on growth at year $t$ are delayed for a specific number of years.

An alternative method that gets around this problem by making current countercyclicality depend essentially upon past observations, is to assume that coefficients follow an AR(1) process. Namely, using the notation from equation 1, for each country $i$ and for each coefficient $j$:

$$a_{jit} = a_{ji,t-1} + \varepsilon_{it}^{a_j}, \varepsilon_{it}^{a_j} \sim N(0, \sigma_{a_j}^2). \tag{4}$$

The main challenge in implementing this method is to estimate $\sigma_{a_j}^2$ (the variance of the coefficients) at the same time as the variance of the observation, i.e. the variance $\sigma_\varepsilon^2$ in the formulation of equation 1. Once these variances are estimated, applying the Kalman filter gives the best estimates for $a_{jit}$.

The optimal estimates for these variance are extremely hard to compute. While finding analytical closed form solutions turns out to be virtually impossible, Markov Chain Monte Carlo (MCMC) methods provide a feasible numerical approximation. We implement the method in Matlab, assuming that the variances of the coefficients and equation are the same for all countries[12]. We are thus left with three variances to estimate: two for the

[12]This assumption is reasonable since the OECD countries in our sample share similar institutions and degrees of economic development. Moreover, this assumption is similar to assuming no heteroskedasiticty across panels when estimating a panel regression, which is the standard assumption. Finally, assuming country-specific variances would

coefficient processes ($\sigma_{a_j}^2, j = 1, 2$) and one for the variance of the error in the equation ($\sigma_\varepsilon^2$). Intuitively, the MCMC method explores randomly (using a Markov chain, hence the name) a wide spectrum of possible values for the variances, and one then retains a set of values that is representative of probable values given the data[13]. An advantage of the MCMC method over maximum likelihood type methods is that it does not get stuck in local solutions and properly represents uncertainty about the variances[14]. Once we obtain the estimates of these three variances, the $a_{jit}$ coefficients can be calculated using the Kalman filter.

AR(1) MCMC is to be preferred over the previous methods for two reasons. First, it reflects a reasonable assumption about policy, i.e. that policy changes slowly and depends on the immediate past. Second, and most importantly, it is econometrically appealing in that it makes policy reflected in the $a_{jit}$ coefficients mainly depend on the past (because of the AR(1) specification)[15]; thus, when the $a_{jit}$ coefficients are used as explanatory variables in panel regressions, it is less likely that there should be a reverse causation problem.

## 2.4 Results

We now use the AR(1) method as described above to characterize the level and time path of the countercyclicality of budget deficits in the OECD countries in our sample. We also report some basic results with the 10

make estimates much more imprecise due to the fact that our relatively small number of observations would have to be used to identify many more parameters.

[13]See appendix 1 for more details on the implementation of this method.

[14]It is indeed also possible to use maximum likelihood type methods to estimate the variances, but these are precisely liable to get stuck in local solutions. In a previous version of this paper, we used such a method, amended so that it does not systematically get stuck in a local solution. In practice, the estimates of the coefficients $a_{jit}$ we had obtained using that method are highly correlated with the ones obtained here using MCMC.

[15]The coefficients also depend on the future in as much as their variance is calculated using the full sample of available observations.

years rolling window and Gaussian weighted OLS methods.

Table 1 summarizes the descriptive statistics of our main variables of interest. For all three measures, the budget deficit is countercyclical (positive coefficient), which is consistent with Lane's (2003) finding that the primary surplus is procyclical. It is worth noting that the three different methods used in the first stage to estimate countercyclicality give very similar results in terms of the mean: a mean of about .5 means that on average in our sample a 1 percentage point increase in the opposite of the GDP gap (i.e. a worse recession) lead to about .5 percentage points increase in the budget deficit as a share of the GDP. In terms of the variance of these measures, we can see that the standard error is largest for the 10-years rolling window method, as expected; it is smaller for the Gaussian method, and even smaller for the AR(1) MCMC method.

*TABLE* 1 *HERE*

We now look at the evolution of the countercyclicality of budget deficit, as measured by the estimated coefficients $a_{1it}$ from equation 1. Figure 2 shows the evolution of the countercyclicality of the budget deficit for the US estimated by the three methods described above. We can readily see that, as expected given the construction of these measures and their empirical standard errors, the 10 years rolling window yields the most volatile results, and the AR(1) method is the smoothest with the Gaussian-weighted OLS method lying in between. Overall, all three methods show an increase in countercyclicality over time, with a recent trend towards decreasing countercyclicality shown by the 10 years rolling window and Gaussian-weighted OLS methods.

*FIGURE* 2 *HERE*

In Figure 3, we then show the countercyclicality of the budget deficit estimated through the AR(1) method for a few countries in our sample. US and UK countercyclicality tends to increase over time, especially since the 1980's. On the contrary, the average countercyclicality of budgetary policy in EMU countries slightly decreases over time. Also, one can observe some divergence between EMU and non-EMU countries: at the beginning of the period, the countercyclicality of the budget deficit in EMU countries was very similar to that in the US, however, as of the 1990's, the US and the UK became significantly more countercyclical whereas the EMU did not.

*FIGURE* 3 *HERE*

In Figure 4, we plot the same evolution, this time based on coefficients that are estimated using the Gaussian-weighted OLS. Trends in estimates are very similar to those obtained using the AR(1) method.

*FIGURE* 4 *HERE*

These results are consistent with Gali and Perotti (2003), who show, splitting their sample by decades, that in general fiscal deficits in the OECD have become more countercyclical, but less so in EMU countries. Here, we confirm these results using a full-fledged time-series measure of countercyclicality.

To summarize our descriptive results, we found that the budget deficit has become more countercyclical in the US and the UK than in EMU countries since the 1990s. In the next section we investigate possible explanations for these observed differences in the countercyclicality of budgetary deficit across countries and over time.

# 3 First stage: determinants of the countercyclicality of budgetary policy

In this section, we use the series of cyclicality coefficients derived using the AR(1) MCMC method and regress the countercyclicality of budgetary policy over a set of macroeconomic variables. Since our sample is restricted to OECD countries, little variation should be expected from the corruption or other institutional variables considered by the literature so far[16]. Instead, we focus on the following candidate variables: financial development, openness, EMU membership[17], and whether the country has adopted inflation targeting. We also include GDP growth volatility as measured by the standard error of GDP growth, lag of log real GDP per capita and the government share of GDP as control variables.

Financial development is a plausible suspect as it influences both the ability and the willingness of governments to borrow in recessions in order to finance the budget deficit. Lower financial development should thus translate into lower countercyclicality of budget deficit. While OECD countries are arguably less subject to borrowing constraints than other countries in the world, there is still a fair amount of cross-country variation in financial development among OECD countries. Openness is also a plausible candidate as one can expect foreign capital to flow in during booms and flow out during recessions, implying that the cost of capital is higher during recessions than during booms. This in turn tends to increase the long-run cost of financing countercyclical budget deficit policies while maintaining

[16]As mentioned above, using ICRG indicators turns out not to be of interest for our analysis.

[17]This dummy variable takes a value of 1 for all countries that currently belong to the EMU, and 0 for all the other countries. This is because the EMU has been prepared for many years so that the countries that would eventually join might be different even before the EMU is fully effective.

the overall debt constant on average over the long run. The EMU dummy is also a plausible candidate, given: (i) our observation in Figures 2 and 3 that the budget deficit is less countercyclical in the Eurozone than in the US or the UK; (ii) the deficit and debt restrictions imposed by the Stability and Growth Pact and also the restrictions that individual countries imposed on themselves in order to qualify for EMU membership.

Inflation targeting should also improve a country's willingness or ability to conduct countercyclical budgetary policy. In particular, one potential factor that might discourage governments to borrow in recessions, is people's expectation that such borrowing might result in higher inflation in the future, for example as a way for the government to partially default on its debt obligations. This in turn would reduce the impact of current government borrowing on private (long-term) investment. Inflation targeting increases the effectiveness of government borrowing in recession by making such expectations less reasonable.

Table 2, where the countercyclicality measures are derived using the AR(1) MCMC method, shows results that are consistent with these conjectures, namely: (i) while countries that are more financially developed tend to have a less countercyclical budgetary policy (column 1), as a country gets more financially developed, it exhibits a significantly more countercyclical budget deficit (column 2); using the results from column 2, our estimates imply that a 10 percentage points increase in private credit over GDP is associated with an increase of about 0.0196 in the countercyclicality of the budget deficit; in other words, it is precisely when the countercyclicality of the budget deficit is more positively correlated with growth, namely when financial development is low, that budgetary deficit countercyclicality seems hardest to achieve; (ii) when using country and year fixed effects (column 2) more trade openness is positively and significantly associated with budgetary deficit countercyclicality (the table shows a positive coefficient on openness); (iii) EMU countries and countries with a larger standard error of GDP growth appear to have a harder time achieving budgetary deficit countercyclicality (column 1) ; the EMU dummy implies that on average EMU countries' budgetary policy countercyclicality is lower by 0.127, which is about a fourth of a standard deviation; the effect of the EMU dummy is more likely to be explained by rigidities already imposed by the precursor EMS regime and then reinforced by the Maastricht Treaty, rather than the 1999 implementation of the EMU itself[18]; further investigation of this question is however beyond the scope of this paper; (iv) a higher share of government in the GDP is associated with a more countercyclical budgetary policy; (v) pursuing inflation targeting is associated with a more countercyclical budget deficit. Note that the coefficient on the inflation targeting dummy in column 2 is of the same magnitude as the coefficient on the EMU dummy in column 1, but of opposite sign.

*TABLE* 2 *HERE*

Hence, a lower level of financial development, a higher degree of openness, belonging to the EMU group, and the absence of inflation targeting, are all associated with a lower degree of countercyclicality in the budget deficit. In the next section we move to second stage analysis of the effect of budget deficit cyclicality on growth.

[18]We have experimented with an interaction between the EMU dummy and a post-1999 dummy, but this interaction was typically insignificant, indicating that there is no substantial change occurring with the full implementation of the EMU in 1999.

# 4 Second stage: cyclical budget deficit and growth

In this section we regress growth on the cyclicality coefficients for budgetary policy derived in Section 2, financial development, the interaction between the two variables, and a set of controls. Our discussion of the theory and microeconomic evidence on volatility, credit constraints and R&D/growth in the Introduction suggests that the lower financial development, the more positive the correlation should be between growth and the countercyclicality of budgetary policy: the idea is that a more countercyclical budgetary policy can help reduce the negative effect that negative liquidity shocks impose on credit-constrained firms that invest in R&D and innovation.

## 4.1 Empirical specification and results

Our empirical specification is:

$$\Delta y_{it} = \beta_1 a_{1i,t-1} + \beta_2 p_{i,t-1} + \beta_3 a_{i,t-1} p_{i,t-1} + X_{it}\beta_4 + \gamma_i + \delta_t + \varepsilon_{it}, \qquad (5)$$

where $\Delta y_{it}$ is the first difference of the log of real GDP per capita in country $i$ and year $t$; $a_{1i,t-1}$ is the countercyclicality of the budget deficit as estimated by the AR(1) MCMC method. Since $a_{1i,t-1}$ is an estimated coefficient, we weigh each observation by the inverse of the variance of this coefficient (aweights in Stata), thus giving higher weight to coefficients that are more precisely estimated in the first stage. $p_{i,t-1}$ is the ratio of private credit to GDP borrowed from Levine (2001); $X_{it}$ is a vector of control variables that vary with the specification considered, $\gamma_i$ is a country fixed effect, $\delta_t$ is a year fixed effect, and $\varepsilon_{it}$ is the error term.

In Table 3, we first report results with a limited set of controls representing the most widely accepted determinants of growth: lag of log real GDP per capita, population growth and investment over GDP (column 1). We then add more controls, namely schooling, trade openness, inflation, government share of GDP and inflation targeting (column 2). Note that since we control for inflation, we indirectly control for the impact of monetary policy on growth.

The prediction is that of a positive $\beta_1$ coefficient for the effect on growth of the countercyclicality of the budget deficit when private credit over GDP is 0, and of a negative $\beta_3$ coefficient on countercyclicality interacted with financial development. In the first column of Table 3, using a limited set of controls, we see that the corresponding coefficients have the anticipated signs and are statistically significant: a more countercyclical budget deficit is positively correlated with growth, but the interaction term between countercyclicality and financial development is negative. Including a richer set of controls in column 2 does not change the results. If anything, the point estimates are larger: a coefficient of 0.11 of the lagged countercyclicality of budget deficit means that if private credit over GDP is 0, then increasing the countercyclicality of the budget deficit by one percentage point increases growth by 0.11 percentage points. For each percentage point increase in private credit over GDP, this positive effect of countercyclicality diminishes by 0.0004. The effect of the interaction is thus small: private credit over GDP would need to be larger than 2.75 for a countercyclical budgetary policy to become growth-reducing. Such a high value of private credit over GDP is not observed in our sample: the US in 2000, at 2.24, has the highest value of this variable in our sample.

Then, in columns 3 and 4, we repeat the same specifications as in columns 1 and 2, but allow the impact of the interaction between the countercyclicality of the budget deficit and private credit over GDP to differ by quartiles of the private credit over GDP (the first quartile is then the excluded category). For example, the dummy "'2ndq (Private Credit/GDP)"' is equal to one if the Private Credit/GDP ratio lies in the second quartile, and is equal to zero otherwise. As the results in these columns show, the interaction between cyclicality and financial development is non-linear, with a significant jump occurring when the private credit ratio moves from the second to the third quartile. In other terms, it is only at fairly high levels of financial development that countercyclical budgetary policy becomes noticeably less growth enhancing.

*TABLE* 3 *HERE*

Table 3 is thus consistent with the prediction of a positive effect of a countercyclical budget deficit on growth, whereas we see a negative and significant interaction effect between private credit and the countercyclicality variable. Thus the less financially developed a country is, the more growth-enhancing it is for the government to be countercyclical in its budgetary policy. In particular, we observe that EMU countries have budgetary policies that are on average far less countercyclical than in the US (0.37 vs. 0.61), even though the US are more financially developed than the EMU: thus, the ratio of private credit to GDP in 2000 in the EMU is equal 1.02 against 2.24 in the US. Then, to the extent that it reflects the causality from cyclical budgetary policy to growth, the regression in Table 3 suggests that increasing the countercyclicality of the budgetary policy would be more growth-enhancing for the EMU than for the US.

## 4.2 Robustness tests

This section discusses various potential issues with our Table 3 estimates. We take as the reference specification for this discussion the specification shown in Table 3, column 2. Therefore, when we report on alternative specifications, they are all based on this reference specification.

A potential first source of concern for our estimation strategy is autocorrelation of residuals, which is typical in panel growth regressions. This implies that the standard errors may be biased. To correct for this potential bias, we used Newey errors to adjust the standard errors in the reference specification. Allowing for autocorrelation of errors up to lag 1 increased the standard errors very slightly and left the coefficients significant at the 1% level. Allowing for autocorrelation up to 5 lags leaves the effect of the countercyclicality of the budget deficit at the same level of statistical significance, but makes the interaction between the countercyclicality of the budget deficit and private credit be only significant at the 2% instead of the 1% level. Globally, it does not seem that autocorrelation of residuals substantially affects the standard errors of our estimates.

Second, the reader may wonder about what components of the budget deficit increase growth when they are more countercyclical. For example, is it the countercyclicality of total government spending that ultimately matters for growth? What about transfers and social security spending? We have run the same analysis for these variables as for the budget deficit[19] and found that their countercyclicality was not significant in explaining economic growth. This indicates that the cyclical behavior of automatic stabilizers is unlikely to fully account for our results: namely, it is not the case that just increasing transfers and social security spending in recessions

[19]Specifically, in equation 1, we replaced the first difference of debt by the first difference of each of these variables.

increases economic growth. What matters for growth is not the countercyclicality of spending per se (be it discretionary or not) but rather the degree to which this spending is financed by debt, i.e. the degree to which the government injects extra liquidity in the economy.

Third, the reader may be interested in knowing what happens if we replace the AR(1) MCMC estimate of countercyclicality by the Gaussian-weighted or the 10-years rolling windows OLS. In the case of the Gaussian, the coefficients on the countercyclicality of the budget deficit and on its interaction with private credit have the same sign as in the reference specification and are significant at the 1% level. The only difference is that the value of the coefficient on the countercyclicality of the budget deficit is lower. In the case of the 10-years rolling windows method, the coefficients of interest are of the same sign, but are not statistically significant, which is not surprising since these estimates are much noisier.

Fourth, one may still be skeptical about the causal interpretation of our estimates. As mentioned in section 2, our AR(1) MCMC estimate of countercyclicality should be in principle mostly uncorrelated with the future, reducing the endogeneity problem. First, to check whether indeed future countercyclicality is independent of growth, we include both the lag and the lead of the countercyclicality measure in the reference specification. Doing so does not significantly change the coefficient on the lagged countercyclicality but yields an insignificant and positive coefficient on the lead of procyclciality. These results are consistent with countercyclicality causing growth and not the reverse. Second, we noticed that inflation targeting is associated with a less countercyclical budget deficit (Table 2) but is insignificant in explaining growth (Table 3). This raises the possibility of using inflation targeting as an instrument for countercyclicality in a GMM framework. In the GMM estimation, we instrument both the countercyclicality variable and the lagged GDP per capita. For the latter, we use the classic instruments second and third lag of GDP per capita. Excluded instruments in our GMM regression are thus second and third lag of GDP per capita and the inflation targeting dummy. Moreover our GMM estimates allow for Newey errors of lag 1. We have thus re-estimated the reference specification using GMM. First stage estimates are significant, but the explanatory power of inflation targeting for countercyclicality is limited. Overidentifying restrictions are not rejected by the J test. However, we do not reject that countercyclicality and its interaction with private credit are exogenous, which means that GMM is not more appropriate than OLS. The coefficients on countercyclicality and its interaction with private credit are of similar magnitudes as in the reference specification but they are not significant (P-values around 30%). This exercise thus confirms that our countercyclicality measure is unlikely to be endogenous.

Finally, one may be interested in the time horizon of our effects: when the countercyclicality of the budget deficit changes in a given year, how far in the future does the effect on growth persist? One way to answer this question is to modify the reference specification by replacing the lag of the countercyclicality of the budget deficit, private credit over GDP and the interaction of the two by further lags. When using the second lag of these variables, the coefficients of interest ($\beta_1$ and $\beta_3$) are still significant and of the same sign, but the $R^2$ diminishes slightly. When using the third lag of these variables, the coefficient on the countercyclicality of the deficit is still significant, but the interaction with private credit is no longer significant. Using even further lags makes the coefficients of interest become insignificant. Thus, it seems that an increase in the countercyclicality of budgetary policy affects GDP growth up to 2 or 3 years later.

# 5 Conclusion

In this paper we have analyzed the dynamics and determinants of the cyclicality of budgetary policy on a yearly panel of OECD countries, and the relationship between this cyclicality, financial development, and economic growth. Our findings can be summarized as follows: first, countercyclicality of budget deficits has generally increased over time. However, in EMU countries, the budget deficit became slightly less countercyclical. Second, countercyclicality of budgetary policy appears to be facilitated by a higher level of financial development, a lower degree of openness to trade, and a monetary policy committed to inflation targeting. Third, we found that countercyclical budget deficits are more positively associated with growth the lower the country's level of financial development.

The line of research pursued in this paper bears potentially interesting growth policy implications. In particular, our second stage regressions suggest that growth in EMU countries could be fostered if the budget deficit in the eurozone became more countercyclical. Our first stage regression suggests that this in turn could be partly achieved by having the EMU area move toward inflation targeting, e.g following the UK lead in this respect, and also by improving the coordination among finance ministers in the eurozone on fiscal policy over the cycle so as to make it become more countercyclical[20].

The analysis in this paper should be seen as one step in a broader research program. First, one could try to perform the same kind of analysis for other groups of countries, e.g middle income countries in Latin America or in Central and Eastern Europe. Second, one could take a similar AABM-type of approach to volatility, financial development and growth

[20]The Sapir report (Sapir et al (2003)) recommended the setting-up of "rainy day" funds supervised by the European Commission.

to further explore the relationship between growth and the conduct of monetary policy. For example, to which extent allowing for higher procyclicality of short term nominal interest rates, can help firms maintain R&D investments in recessions and/or improve governments' ability to implement growth-enhancing countercyclical budgetary policies? Finally, one could investigate the possible interactions in growth regressions between countercyclical budgetary policy and structural reforms in the product and labor markets.

# References


[1] Acemoglu, D, Johnson, S, Robinson, J, and Y. Thaicharoen (2003), "Institutional Causes, Macroeconomic Symptoms: Volatility, Crises, and Growth", *Journal of Monetary Economics,* 50,49-123.

[2] Aghion, P, Angeletos, M, Banerjee, A, and K. Manova (2005), "Volatility and Growth: Credit Constraints and Productivity-Enhancing Investment", NBER Working Paper No 11349.

[3] Aghion, P, Bacchetta, P, Ranciere, R, and K. Rogoff (2006), "Exchange Rate Volatility and Productivity Growth: The Role of Financial Development", NBER Working Paper No 12117.

[4] Alesina, A, and R. Perotti. (1996), "Fiscal Adjustments in OECD Countries: Composition and Macroeconomic Effects", NBER Working Paper No 5730.

[5] Alesina, A, and G. Tabellini (2005), "Why is Fiscal Policy often Procyclical?", NBER Working Paper No 11600.

[6] Atkeson, A, and C. Phelan (1994), " Reconsidering the Costs of Business Cycles with Incomplete Markets", NBER Macroeconomics Annual, 9, 187-207.

[7] Berman, N, Eymard, L, Aghion, P, Askenazy, P, and G. Cette (2007), "Credit Constraints and Cyclical R&D Investment: Evidence from French Firm-Level Panel Data", mimeo Banque de France.

[8] Barro, R (1979), "On the Determination of Public Debt", *Journal of Political Economy,* 87, pp. 940-971.

[9] Batalgi, B (2001), *Econometric Analysis of Panel Data,* 2d ed. New York: John Wiley & Sons.

[10] Berman, N, Eymard, L, Aghion, P, Askenazy, P, and G. Cette (2007), "Credit Constraints and the Cyclicality of R&D Investment: Evidence from French Firm-Level Panel Data", working paper.

[11] Calderon, C, Duncan, R, and K. Schmidt-Hebbel (2004), "Institutions and Cyclical Properties of Macroeconomic Policies", Central Bank of Chile Working Papers No 285.

[12] Chib,S, and E. Greenberg (1995), "Understanding the Metropolis-Hastings Algorithm", *The American Statistician*, 49, 327-335.

[13] Easterly, W, (2005), "National Policies and Economic Growth: A Reappraisal", Chapter 15 in *Handbook of Economic Growth,* P. Aghion and S. Durlauf eds.

[14] Gali, J, and R. Perotti (2003), "Fiscal Policy and Monetary Integration in Europe", *Economic Policy*, 533-572.

[15] Hallerberg M., Strauch R., von Hagen J. (2004), "The design of fiscal rules and forms of governance in European Union countries," Working Paper Series 419, European Central Bank.

[16] Kording K. and Marinescu I. (2006), "Sampling approaches to Kalman learning", work in progress.

[17] Lane, P, (2003), "The Cyclical Behavior of Fiscal Policy: Evidence from the OECD", *Journal of Public Economics,* 87, 2661-2675.

[18] Lane, P, and A. Tornell (1998), "Why Aren't Latin American Savings Rates Procyclical?, *Journal of Development Economics*, 57, 185-200.

[19] Lucas, R (1987), *Models of Business Cycles,* New-York: Basil Blackwell.

[20] OECD (1995), "Estimating Potential Output, Output Gaps and Structural Budget Balances", by Claude Giorno, Pete Richardson, Deborah Roseveare and Paul van den Noord, OECD Economics Department Working Paper No. 152.

[21] Talvi, E, and C. Vegh (2000), "Tax Base Variability and Procyclical Fiscal Policy", NBER Working Paper No 7499.

[22] Vega M. and Winkelried D. (2005), "Inflation targeting and inflation behavior: a successful story?", International Journal of Central Banking, Vol. 1, No. 3, December 2005, 153-175.

**Table 1: Summary statistics**

| | Obs. | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| **GDP gap** | 756 | 0.000 | 0.019 | -0.070 | 0.071 |
| **Gross government debt/GDP** | 756 | 0.548 | 0.295 | 0.046 | 1.686 |
| **Budget deficit/GDP** | 756 | 0.048 | 0.046 | -0.065 | 0.321 |
| **Countercyclicality of budget deficit (AR(1))** | 641 | 0.511 | 0.563 | 2.686 | -0.342 |
| **Countercyclicality of budget deficit (Gaussian weighted rolling window)** | 756 | 0.578 | 0.752 | 3.337 | -1.112 |
| **Countercyclicality of budget deficit (10-years rolling window)** | 532 | 0.608 | 1.065 | 8.972 | -3.011 |
| **Growth of GDP per capita** | 689 | 0.025 | 0.026 | -0.092 | 0.116 |
| **Private credit/GDP** | 585 | 0.801 | 0.392 | 0.128 | 2.240 |
| **Average years of schooling for the population over 25 years old** | 585 | 8.236 | 1.989 | 2.510 | 12.250 |
| **Openness** | 605 | 53.633 | 35.641 | 8.705 | 266.883 |
| **Inflation** | 756 | 0.061 | 0.066 | -0.025 | 0.762 |
| **Population growth** | 689 | 0.006 | 0.005 | -0.018 | 0.047 |
| **Government share of GDP (in %)** | 605 | 12.440 | 5.709 | 3.008 | 27.848 |
| **Investment/GDP (in%)** | 605 | 24.106 | 4.983 | 12.867 | 41.635 |
| **Inflation targeting dummy** | 756 | 0.112 | 0.316 | 0 | 1 |

Note: sample restricted to observations where the countercyclicality of budget deficit computed using Gaussian weighted rolling windows is not missing.
Source: OECD Economic Outlook, Levine dataset, Barro Lee dataset, Penn World Tables 6.1.

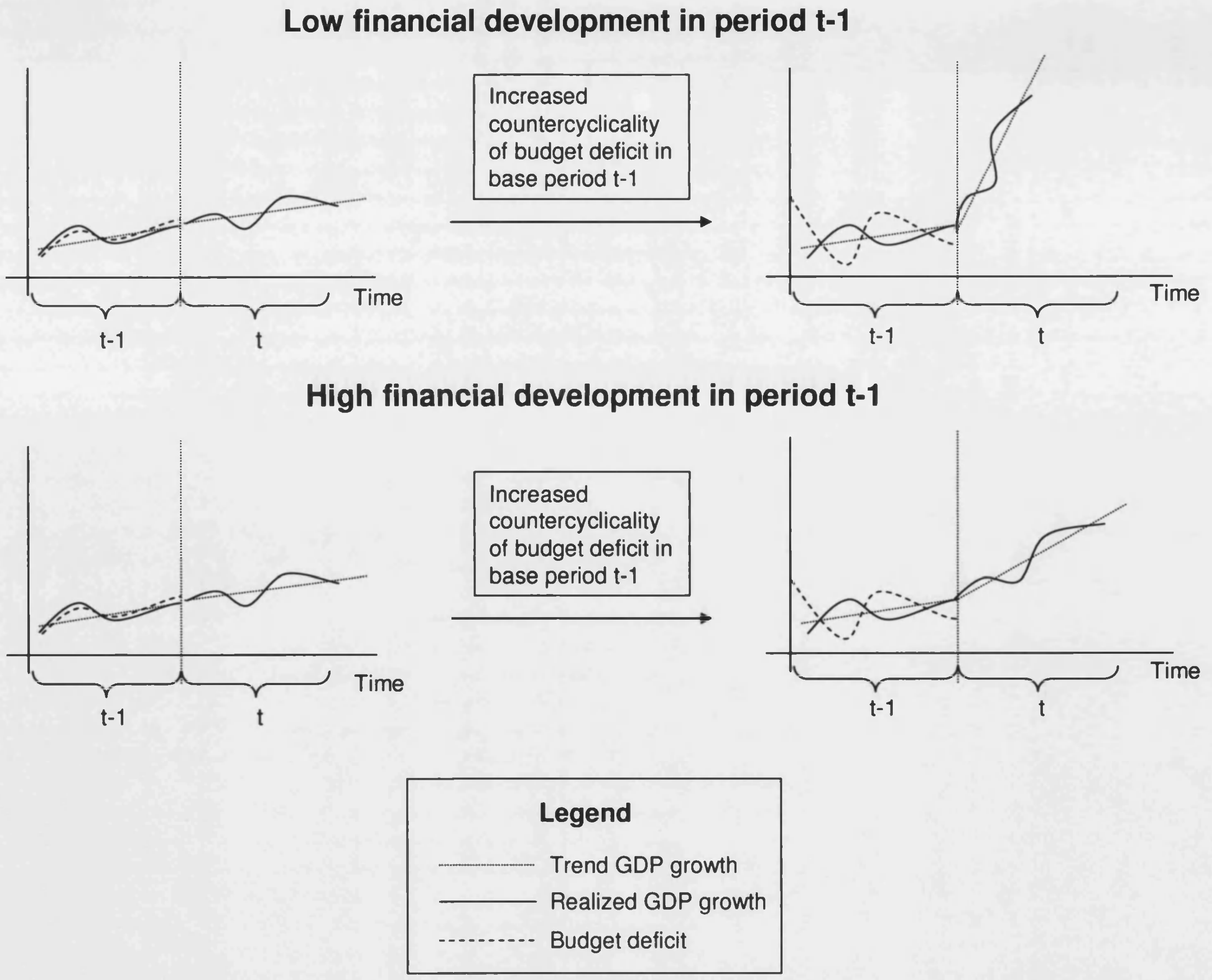
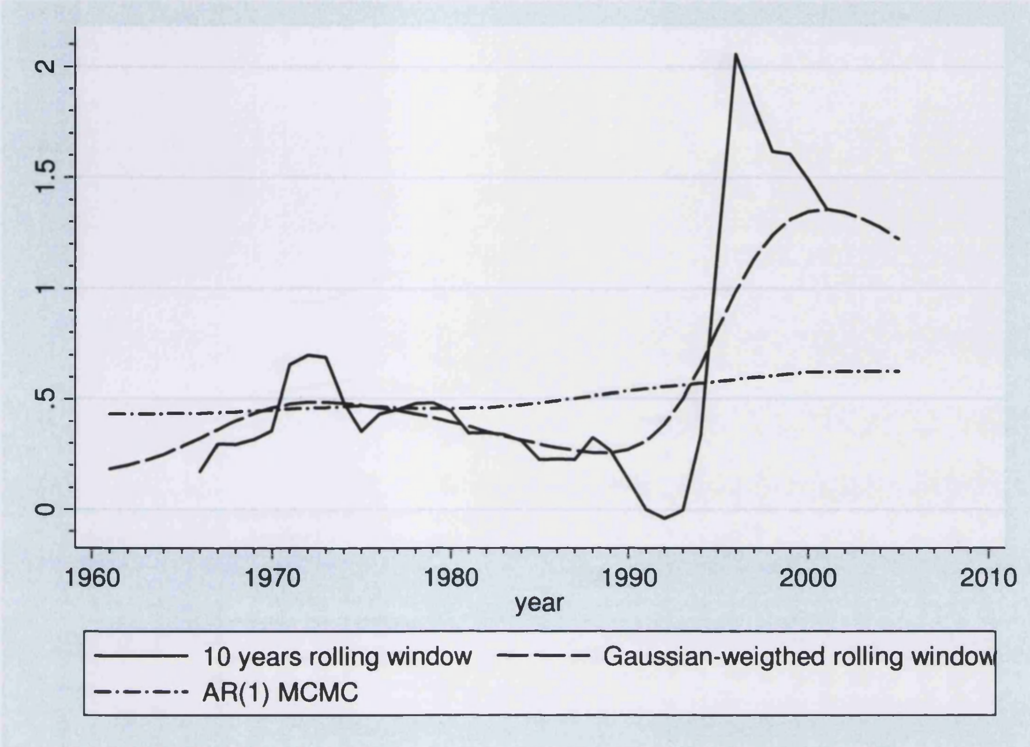


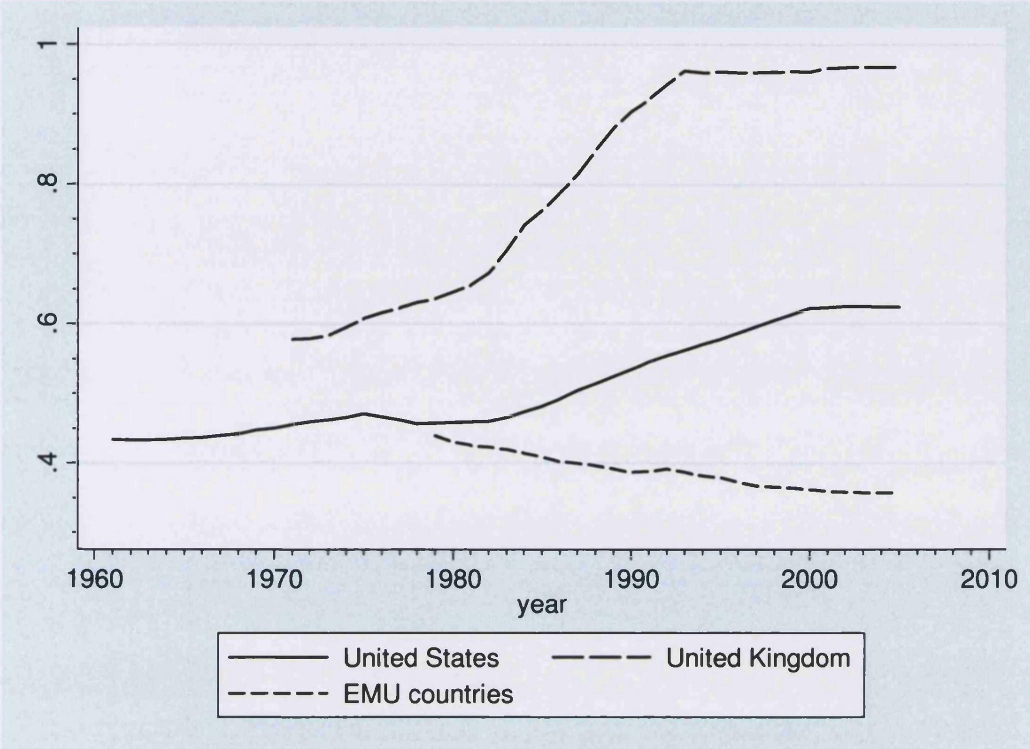

**Figure 1: The impact of an increase in the countercyclicality of the budget deficit on growth**

**Figure 2: the countercyclicality of the budget deficit in the USA**



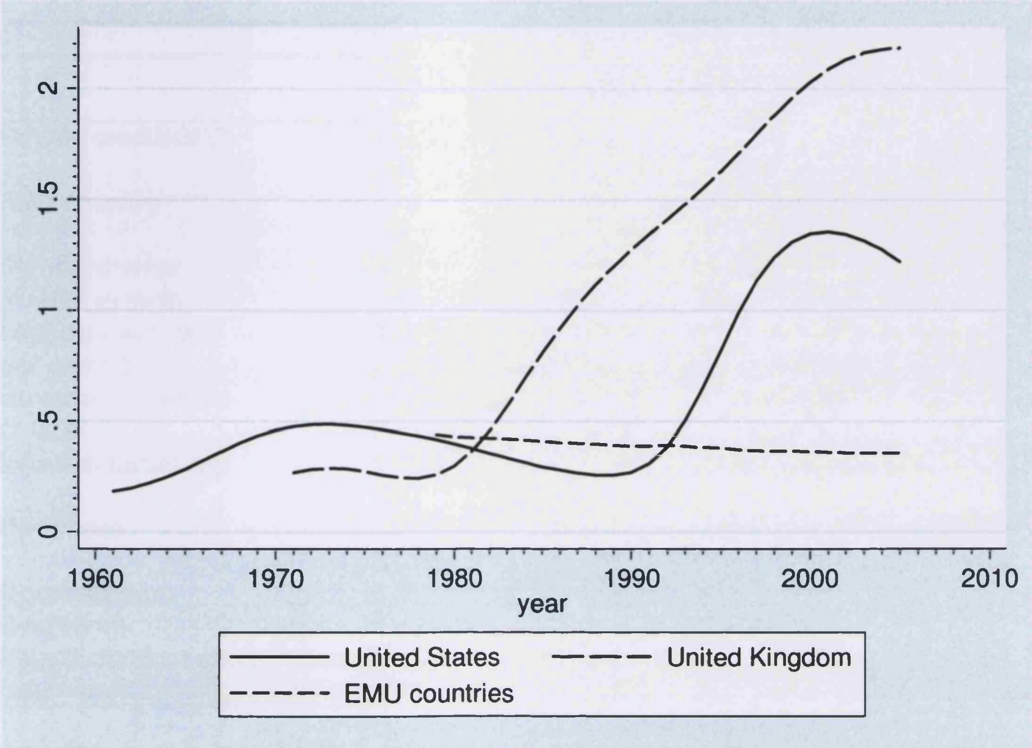

Note: the graph plots the $a_{1it}$ coefficients, i.e. the coefficients on the opposite of the output gap in equation 1, using various estimation techniques.
Source: OECD Economic Outlook.

**Figure 3: The countercyclicality of budget deficits using the AR(1) MCMC method**




Note: the graph plots the $a_{1it}$ coefficients, i.e. the coefficients on the opposite of the output gap in equation 1, using the AR(1) MCMC method. For EMU countries (i.e. countries who are or will be part of the EMU), the line represents the average of the estimated coefficients for the EMU countries present in the sample; the average is only computed for those years where all EMU countries have non-missing observations.
Source: OECD Economic Outlook.

**Figure 4: The countercyclicality of budget deficits using the Gaussian-weighted OLS method**




Note: the graph plots the $a_{1it}$ coefficients, i.e. the coefficients on the opposite of the output gap in equation 1, using the Gaussian-weighted rolling window OLS method. For EMU countries (i.e. countries who are or will be part of the EMU), the line represents the average of the estimated coefficients for the EMU countries present in the sample; the average is only computed for those years where all EMU countries have non-missing observations.
Source: OECD Economic Outlook.

**Table 2: The determinants of the countercyclicality of budget deficits**

| | (1) | (2) |
|---|---|---|
| | Year f.e. | Country year f.e. |
| **Private credit/GDP** | -0.453 | 0.196 |
| | (0.115)*** | (0.018)*** |
| **EMU country** | -0.127 | |
| | (0.038)*** | |
| **Standard error of GDP growth** | -3.364 | |
| | (0.818)*** | |
| **Lag(log (real GDP per capita))** | 0.011 | 0.072 |
| | (0.017) | (0.071) |
| **Government share of GDP (in %)** | 0.000 | 0.004 |
| | (0.005) | (0.001)*** |
| **Inflation targeting** | 0.292 | 0.112 |
| | (0.081)*** | (0.015)*** |
| **Openness** | -0.007 | -0.002 |
| | (0.001)*** | (0.001)*** |
| **Observations** | 515 | 515 |
| **R-squared** | 0.21 | 0.99 |

Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Note: The explained variable is the coefficient on the opposite of the GDP gap in equation 1, estimated using the AR(1) MCMC method. EMU country is a dummy variable equal to 1 for all countries that are part of the EMU as of 2006.
Source: OECD Economic Outlook, Levine dataset, Barro Lee dataset, Penn World Tables 6.1.

**Table 3: The effect of the countercyclicality of budget deficits on growth, AR(1) MCMC method**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| **lag(Countercyclicality of budget deficit)** | 0.075 | 0.110 | 0.058 | 0.081 |
| | (0.021)*** | (0.024)*** | (0.016)*** | (0.018)*** |
| **lag(Private credit/GDP)** | -0.010 | -0.005 | -0.014 | -0.008 |
| | (0.008) | (0.008) | (0.007)** | (0.007) |
| **lag(Countercyclicality of budget deficit*Private credit/GDP)** | -0.030 | -0.040 | | |
| | (0.012)*** | (0.014)*** | | |
| **lag(Countercyclicality of budget deficit*2ndq(Private credit/GDP))** | | | -0.006 | -0.009 |
| | | | (0.003)** | (0.003)*** |
| **lag(Countercyclicality of budget deficit*3rdq(Private credit/GDP))** | | | -0.022 | -0.024 |
| | | | (0.007)*** | (0.008)*** |
| **lag(Countercyclicality of budget deficit*4thq(Private credit/GDP))** | | | -0.023 | -0.030 |
| | | | (0.008)*** | (0.010)*** |
| **lag(log (real GDP per capita))** | -0.140 | -0.132 | -0.142 | -0.131 |
| | (0.022)*** | (0.022)*** | (0.022)*** | (0.022)*** |
| **Investment/GDP (in%)** | 0.002 | 0.002 | 0.002 | 0.002 |
| | (0.000)*** | (0.000)*** | (0.000)*** | (0.000)*** |
| **Population growth** | -1.490 | -1.702 | -1.484 | -1.635 |
| | (0.268)*** | (0.284)*** | (0.272)*** | (0.290)*** |
| **Average years of schooling for the population over 25 years old** | | 0.002 | | 0.003 |
| | | (0.003) | | (0.003) |
| **Government share of GDP (in %)** | | -0.001 | | -0.001 |
| | | (0.000)*** | | (0.000)** |
| **Inflation** | | -0.049 | | -0.053 |
| | | (0.022)** | | (0.021)** |
| **Inflation targeting** | | -0.004 | | -0.001 |
| | | (0.005) | | (0.005) |
| **Openness** | | 0.001 | | 0.001 |
| | | (0.000)** | | (0.000)*** |
| **Observations** | 477 | 467 | 477 | 467 |
| **R-squared** | 0.60 | 0.64 | 0.62 | 0.65 |

Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Note: The explained variable is the first difference of the log of real GDP per capita. All specifications include country and year fixed effects. Columns 3 and 4 allow for the effects of countercyclicality of the budget deficit to differ with quartiles of private credit/GDP.
Source: OECD Economic Outlook, Levine dataset, Barro Lee dataset, Penn World Tables 6.1.

# Appendix 1: the AR(1) MCMC method for calculating cyclicality in the first stage

The aim of this section is to give a brief description of how we used the Kalman filter together with Markov Chain Monte Carlo methods (MCMC) in order to estimate the coefficients $a_{jit}$ from equation 1 under the assumption that they follow an AR(1) process as desribed by equation 4. The implementation was carried out in Matlab.

Estimating the means and variances of the coefficients of interest - that is $a_{jit}$ in equation 4 - involves two procedures: Kalman filtering[1] and MCMC.

To compute the coefficients with the Kalman filter for each country, we need to know the values of three variances :

- $\sigma^2_{a_j}$ in equation 4, for $j = 1, 2$, i.e. the process variances in the terminology of the Kalman filter.
- the variance $\sigma^2_\varepsilon$ of the error term $\varepsilon_t$ in equation 1, i.e. the measurement error variance in the terminology of the Kalman filter.

Moreover, to use the Kalman filter, we need a prior for the first period of observation for each country, that is a specification of our expectation over the values $a_{jit}$ at the first time step. As we do not have any meaningful prior information about cyclicality at the first observed period, we use a very high variance around the prior mean, so that this prior has a negligible effect on the estimates. Specifically, the set of initial values for the coefficients were chosen to be the OLS estimates of the coefficients using the first 10 years

[1]For an excellent overview of the Kalman filter and smoother, see the notes by Max Welling "Kalman Filters", available on the web at http://www.ics.uci.edu/~welling/classnotes/classnotes.html.

of data for each country, and the value of the intial variance is set to be 100000 times the estimated variance of these coefficients.

However, the process variances $\sigma^2_{a_j}$ and the measurement error variance $\sigma^2_\varepsilon$ are unknown and we do not have any meaningful prior over them. We therefore need a method to find reasonable values for these three unknown variances. This is where MCMC methods are useful .

One can think of MCMC as the opposite of simulating. In the case of simulation we know the parameters of our process, for example the variances, and every time we run a simulation program, it gives us a set of possible observed data. More specifically, the probability of getting any set of observed data is the probability defined by the model that we have and the parameters. MCMC is the opposite: we assume that we have a given dataset, and we are producing a set of possible parameters. This is done in such a fashion that the probability of accepting a parameter value is identical to the probability that this parameter value has actually produced the data.

Specifically, in our implementation, we use the classic Metropolis-Hastings (MH) sampler to do MCMC (for an introduction to MCMC and Metropolis-Hastings, see for example Chib and Greenberg (1995)). In MH one starts with arbitrary parameters values. Every iteration one proposes a random change (in our case a small gaussian change) of the parameters. This is what is called the proposal distribution. Subsequently, this change is either accepted or rejected. The probability of acceptance is:

$$p_{accept} = min\left(1, \frac{p(data|new\_parameters)}{p(data|previous\_parameters)}\right) \tag{1}$$

It is easy to prove that this procedure is actually sampling from the correct posterior distribution over the parameter values.

MCMC algorithms go through two different stages. In the first stage the sampler converges to a probable interpretation of the data in terms of the parameters. This stage is called burn-in and took about 500 iterations in our case. Within these 500 iterations, probabilities increased dramatically and then converged to a stable high level. Afterwards, the MCMC algorithm is exploring the space of relevant parameters. Over 3 runs, we took 10000 samples per run after the end of burn-in. To avoid the autocorrelation that typically characterizes a Markov Chain, we only retain samples every 100 iterations in order to compute the final estimates. From these 3 runs, we thus get a total of 300 essentially uncorrelated samples for each of the three parameters we wish to estimate. Convergence of the Markov chain was assessed comparing the within chain correlation with the across chain correlation. From these 300 samples, we can then directly estimate means and variances of the three parameters of interest.

In order to correctly infer the effect of cyclicality on growth in our second stage regressions, we need to determine not only the value of the cyclicality ($a_{1it}$), but also the uncertainty we have about it. To estimate this uncertainty, or in other words the standard deviation of the cyclicality estimates, it is necessary to consider the relevant sources of uncertainty. Two sources are relevant in our case. One is the uncertainty that is represented by the Kalman filter that stems from the finite number of noisy observations. The other source of uncertainty is uncertainty about the three parameters that are modeled by the MCMC process. To combine them, we use the approximation $variance_{total} = variance_{MCMC} + \overline{variance_{Kalman}}$, where $\overline{variance_{Kalman}}$ denotes the average variance over the 300 Kalman filter runs using the 300 samples that we retained from the MCMC estimates of the three variances. This approximation becomes correct if the variance as estimated by the Kalman filter is similar over different runs of the Markov chain, which was a good approximation for our data.

Finally, a full general statistical description of the methods used here can be found in Kording-Marinescu(2006).

# Conclusion

Promoting productivity, job security, and investment are all essential objectives of public policy. In this thesis, I have provided new evidence on the economic impact of firing costs and countercyclical macroeconomic policy.

In the first chapter, I investigated the impact of a reduction in the length of the probationary period in the UK from two years to one year in 1999. I started with a simple model of firms' behavior, where firms learn about the quality of their match with a given worker through signals of the worker's performance. Firms decide to fire depending on firing costs and their belief about match quality. The model shows that firms fire more right before the end of the probationary period and less afterwards; moreover, the shortening of the probationary period form two to one year entails a lower firing hazard for workers with one to two years tenure. Firm may also react to the shortening of the probationary period by recruiting better workers or monitoring more efficiently the workers they already have. The model allows me to predict the effect of such reactions on the firing hazard. The empirical analysis shows a 30% decrease in the firing hazard for workers with one to two years tenure, as predicted by the model. Moreover, the hazard of firing for workers with zero to one year tenure decreases by about 30% as well. This latter result is consistent with firms having increased recruitment quality after the reform. The calibration of the model shows that recruitment quality indeed increased most after the reform, but monitoring intensity also increased somewhat. Finally, there is no evidence of the reform having had a negative impact on employment or unemployment. Thus, the decrease in the probationary period likely increased productivity, with limited or no negative impact on unemployment.

In the second chapter, I generalize the model of firm's firing decision used in the first chapter. First, I note that the firm's situation in the firing context generalizes to all situations when a person or organization is in a relationship and has to decide whether to continue or separate from the relationship. In chapter 2, I develop in more detail the logic of the simple model used in chapter 1, and I analyze more general cases. Thus, in chapter 1, I had assumed that match quality can only take two values. In chapter 2, I examine the case when match quality is normally distributed. I also allow match quality to vary over time, following an AR(1) process; in this case, I use the Kalman filter to solve the agent's problem. I show that most of the basic conclusions from the simple two-quality model carry over to more general assumptions. In particular, the effect of firing costs (or, more generally, separation costs), is particularly insensitive to these assumptions: in all cases, higher firing costs decrease the agent's willingness to fire at all relationship lengths, and hence decrease the hazard of separation. By contrast, the effect on the separation hazard of particular sources of uncertainty does depend on whether match quality is allowed to vary over time; differences are very small at short tenures but become bigger at longer tenures. This is explained by the dynamics of the evolution of match quality over time. If match quality is constant over time, then after a while, all low quality relationships have been terminated, and only high quality relationships remain, which drives the hazard to 0, irrespective of the value of uncertainty parameters[1]. If, on the other hand, match quality can change over time, then a good relationship always has a positive probability of turning bad in the future, and so the separation hazard never declines to 0. The model developed in chapter 2 is useful to analyze relationships such as the employment relationship, marriage or firm-supplier relationships. The model allows us to understand which factors drive the separation hazard and how, and such knowledge is useful when analyzing the impact of some variable (such as the firing cost) on the separation hazard. Moreover, the model can be used for structural estimation, whereby underlying parameters can be estimated[2].

In chapters 1 and 2, I analyzed the impact of firing costs generated by the workers' right to claim unfair dismissal. In these chapters, I implicitly assumed that these costs are fixed and known ex ante. This may be a reasonable approximation, but in practice the actual costs of firing depend on how judges apply the law when a worker sues for unfair dismissal. In chapter 3, I thus analyze the determinants of judge's decisions in unfair dismissal cases, concentrating on the effect of economic conditions. The idea is that if unemployment rates are higher, then it is harder for workers to find a job and thus judges may be inclined to decide in their favor. On the other hand, higher unemployment rates also usually go hand in hand with worse market conditions for firms and judges may be more willing to listen to firms' arguments. The question is then which party judges' decisions favor when economic conditions are worse. The empirical analysis of the effect of

[1] This is only true as long as uncertainty is not total, of course.

[2] Obviously, the more we know about the parameters, the more precise is the estimation. If we don't know anything but the shape of the separation hazard, it is not possible to disentangle the effect of some of the parameters.

economic conditions on judges' decisions is complicated by the fact that cases reaching trial are a selected sample of all possible cases. However, the nature of my dataset allows me to control very well for such selection biases. I find that a higher unemployment or bankruptcy rate induces judges to decide more often in favor of firms, thus rejecting workers' demands. However, if the worker is still unemployed at the moment of the trial, then a higher unemployment rate makes judges more likely to decide in the worker's favor. Judges thus seem to compromise between firms' and workers' interests. It may be that judges consider that if the worker has found a new job, then it is more important to avoid firms having to pay firing costs, since such financial burden may increase the risk of bankruptcy and further job destruction. However, the action of judges can only have a limited impact on bankruptcy, because their decisions do not directly contribute to fostering economic recovery.

Traditionally, macroeconomic policy has been assigned the role of stabilizing the economy over the cycle, thus promoting economic growth. In recent years, however, it has been argued that macroeconomic policy may not be able to increase long run growth, the latter mainly determined by institutions. In chapter 4, Philippe Aghion and I argue that countercyclical debt policy can and does increase economic growth. The theoretical argument is as follows. During recessions, the opportunity cost of investing in new technologies and products decreases because the returns to simply increasing production diminish. However, firms are often credit constrained and thus cannot borrow enough to engage in those growth-enhancing investments. We argue that the government can alleviate firms' credit constraints by borrowing and spending more in bad times, and repaying its debt in good times. This positive effect of government intervention should be stronger the more widespread credit constraints are; in other terms, the effect should be stronger the lower the level of financial development. Empirically, we use a panel dataset of OECD countries to test whether countercyclical debt policy is more growth enhancing when private credit is less abundant. In order to test this hypothesis, we first estimate a time-series of public debt countercyclicality using a series of different methods. Our preferred method assumes that government policy follows and AR(1) process, i.e. that it changes over time but only slowly; as in chapter 2, the use of Kalman filtering techniques is necessary to solve the problem. We find that indeed countercyclical debt policy is more growth-enhancing when private credit over GDP is lower.

The findings presented in this thesis contribute to essential contemporary debates about public policy. In particular, they contribute to explain the difference in the economic performance of European countries versus the United States. Indeed, firing costs have often been blamed for less flexible labor markets and higher unemployment in Europe, the so-called Eurosclerosis. The findings presented in this thesis contribute to the literature establishing that this explanation is probably misleading: firing costs do not necessarily have big negative effects on employment, and they may even contribute to increasing productivity. On the other hand, part of the growth differential between the Eurozone and the US may be explained by the fact that the Eurozone has a less countercyclical debt policy, even though it is less financially developed than the US. The Eurozone may thus increase its growth by becoming more countercyclical.

The papers forming this thesis analyze specific economic problems, starting from clear theoretical premises, and using appropriate statistical tools. They make a contribution to our knowledge of the economic impact of firing costs and countercyclical macroeconomic policy. They also develop useful methods which should lead to fruitful new empirical research. For example, one could use the model from chapter 2 to infer how heterogeneous the population of potential matches is and thus shed some light on job transitions, wage changes and the tenure-wage profile. Chapter 4 has analyzed the impact of countercyclical debt policy on economic growth within the OECD. Using the same technique to compute countercyclicality, one could analyze the effect of other macro policies. Moreover, it would be interesting to see if the effects found for the OECD carry over to developing countries, where credit constraint problems are more severe. These are but some of the questions that could be investigated using insights from the work presented here.