Essays on Labor Market Policies Evaluation

Thomas Le Barbanchon

École Polytechnique

THÈSE

Pour obtenir le grade de
Docteur de l'École Polytechnique
en Sciences Économiques

Présentée et soutenue publiquement le 18 juin 2012 par

Thomas Le Barbanchon

Essays on Labor Market Policies Evaluation

Directeur de thèse : M. Pierre Cahuc

Composition du jury :

Directeur :	M. Pierre Cahuc	-	Professeur à l'Ecole Polytechnique
Rapporteurs :	M. Denis Fougère	-	Directeur de recherche au CNRS
	M. Rafael Lalive	-	Professeur à HEC Lausanne
Suffragants :	M. Christian Belzil	-	Professeur à l'Ecole Polytechnique
	M. Imran Rasul	-	Professeur à University College London

Acknowledgement

I thank Pierre Cahuc for his great support and supervision.

I also thank my coauthors, Luc Behaghel, Bruno Crépon, Thomas Deroyon and Marc Gurgand. The different chapters of my thesis have benefited from various helps along the way and comments in seminars and conferences.

Chapter 1: I thank seminar participants in SOLE 2012, IZA, Crest, J-Pal Europe, INED and Centre d''Etudes pour l'Emploi for their helpful comments. We especially thank the French public employment service, Pole Emploi, for its commitment in implementing this experiment; in particular, I thank François Aventur, Camille Bouchardeau, Danielle Greco, Odile Marchal and Dominique Vernaudon-Prat for their constant support. I thank Andrea Lepine, Julie Moschion and Pascal Achard for excellent research assistance, Abderrazak Chebira, Yann Algan, Corinne Prost and Hélene Garner for very helpful discussion on the measurement of applicants' origin.

Chapter 2: I thank participants at the 2011 Royal Economic Society Annual conference, at the 2011 European Association of Labor Economist annual conference and at seminars in the Crest, and especially Luc Behaghel, Pierre Cahuc, Bruno Crépon, Marc Gurgand, Thierry Kamionka, Julie Labarthe, Guy Laroque, Etienne Lehmann, Cyril Nouveau, Roland Rathelot, Konstantinos Tatsiramos and Petra Todd for helpful remarks. I thank the DARES (Ministry of Labor), the French Employment Agency (Pole emploi) and the French Statistical Institute (Insee) for providing the data.

Chapter 3: I thank Laurent Davezies, Esther Duflo and seminar participants at CREST for their comments on previous versions of this chapter. I also thank the French public employment service (ANPE and Unédic) and the DARES (Ministry of Labir) for their involvement in the experiment from which the data of the application are drawn, and for their financial support to the evaluation.

Chapter 4: I thank seminars participants at the 2009 Tallin EALE conference, at the 2009 EEA conference and at the 2010 TEPP summer school for their helpful comments. I thank Etienne Lehmann, Susan Vroman and two anonymous referees for helpful remarks.

Last, but not least. I thank Caroline for her everyday support.

Contents

Chapter 1

General Introduction

Contents

Labor market policies represent a significant share of Gross Domestic Product. In 2009, OECD countries spent 1.7% of their GDP on labor market policies (this amounts to 2.4% in France)[1]. These policies aim at improving resources allocation in the labor market, which in turn should improve allocation in the whole economy. Broadly speaking, labor market policies aim at decreasing the unemployment rate. For example, Job Search Assistance programs aim at accelerating transitions from unemployment to employment. However, the unemployment rate is not always a sufficient indicator to address all types of resources misallocation. For example, taste-based discrimination in hiring may put to work low productivity majority workers at the expense of high productivity minority workers. Then anti discriminatory labor market policies can succeed even if the global unemployment rate stays flat, as long as they reduce differential treatment, or differences in unemployment rates.

Pointing to high resource misallocation, the French unemployment rate is structurally high compared to other OECD countries (8% vs. 5.8% in 2007 before the Great Recession). The French unemployment rate is around 3 points lower during the 2000s than during the 1990s (see figure 1).

In the French context, commonly cited causes for high unemployment and resource misallocation on the labor market are skill mismatch, search frictions, generous unemployment benefits, distortive and high labor tax combined with high minimum wage, rigid wage bargaining and other rigid labor institutions (see Cahuc and Postel-Vinay (2002) or Blanchard and Landier (2002)).

[1] In 2009, France spent 27 billions euros on passive labor market policies (LMP) with 1.6 million participants and 19 billions euros on active LMP with 70 000 participants. Source: OECD.

Figure 1.1: Unemployment rate in France

Source: OECD stats.

To alleviate negative consequences of high statutory minimum wage, employer labor tax rate has been reduced at the lower part of the wage distribution at several occasions in the 90s and 2000s. This tax reduction policy decreases labor cost by 26% (since 2003) for workers paid at the minimum wage level, boosting labor demand for those workers. Several studies (for example Crépon and Desplatz (2001) and Blanchet, Barlet, and Barbanchon (2009)) conclude to positive employment effects. Costs associated to this permanent tax reduction represent 1.5% of GDP, making that policy the most expensive labor market policy in France. In this thesis, we consider Job Search Assistance policies[2] reducing search frictions and passive labor market policies (as unemployment insurance) which could increase match quality. Search frictions and low quality matching typically arise because

[2]Job Search Assistance belongs to the broader category of Active Labor Market Policies which also encompass training programs and employment creation schemes (wage subsidy in the private sector or direct job creation in the public sector). While training is usually a long process through which new productive skills are acquired, Job Search Assistance programs are shorter programs and aim at monitoring search effort and increasing job search efficiency. JSA enhances job search efficiency by proposing skills assessment exercises, application and job search courses. Assistance can be delivered in collective format (workshops, job clubs) but also during regular individual interviews with job counselor. During those interviews, the counselor proposes job offers and both job counselor and unemployed consider the application strategy to adopt. During such interviews, job search is monitored and unemployed can even been sanctioned when job search is evaluated too low. Apart from this sanctioning dimension, Job Search Assistance typically reduces information imperfection on the labor market.

of imperfect information and congestion externalities. We also consider anti discriminatory labor market policies, such as anonymous applications, which aims at alleviating negative consequences of imperfect information in the matching process.

More precisely, in the first two chapters, we estimate the impact of two labor market policies in the French context:

- anonymous applications: information, such as name, gender, age, nationality and address, is erased from resumes sent to employers (first chapter). Anonymous applications prevent differential treatment based on gender, age...
- Unemployment Insurance generosity: job losers receive benefits to prevent revenue loss and to subsidize job search (second chapter)

We devote special attention to identify causal impacts and thus rely on experimental or quasi experimental evidence to perform our empirical evaluation exercises. Namely, we compare treated and control groups which are ex ante statistically identical. However, we discuss two potential caveats of randomized experiments, one "practical" in the third chapter and one "theoretical" in the fourth chapter:

- ex post comparability of control and treated groups when sample attrition can be different among experimental groups (third chapter)
- uncontrolled contamination between control and treated group through market interaction (fourth chapter)

We discuss those caveats in the context of Job Search Assistance evaluation.

1.1 Empirical analysis of two labor market policies

1.1.1 Anonymous applications

While unemployment insurance and job search counseling are well established and wide spread policies (respectively 1.4% and 1% of GDP in France), anonymous applications have only recently arrived on the public policy stage. Anonymous resumes are used in Belgium in the federal administration. They have been tested locally in the Netherlands, Sweden and in Switzerland, but the tests have not led to new legislation. In Great Britain, an attempt by Liberal Democrats to impose anonymous resumes for recruitment has been opposed by the government in 2009. In France, which we study here, a law was passed in 2006 that made the use of anonymous resumes mandatory for firms with more than 50 employees; yet, the government did not take the steps to define the conditions under which the law would apply.

Anonymous applications aim at reducing pervasive differential treatment on the labor market. Differential treatment by race and gender are repeatedly documented as a prominent feature in many labor markets. Correspondence testing studies, initially primarily developed in the UK, have been instrumental in providing direct and compelling evidence of differential treatment in interview rates. A notable example for race in the US is Bertrand and Mullainathan (2004). Correspondence studies have been increasingly used in France, the country under study here: see in particular Duguet, L´ Horty, Meurs, and Petit (2010) introducing the special issue of *Annals of Economics and Statistics*, n˚99/100, on measuring discrimination. The exact interpretation of differential treatment in terms of discrimination is still debated. In this thesis, we abstract from this normative and important question and focus in the context of a randomized experiment on the actual effectiveness of anonymous application to reduce differential treatment in access to interviews and hires. More precisely, we consider differential treatments between women and men, and between majority workers and workers with foreign background[3] or living in deprived neighborhood.

Given correspondence testing evidence, there is a consensus on the likely effectiveness of anonymous applications to change the pool of candidates called for an interview. However effects on later stages of the hiring process, when candidates are actually interviewed, are controversial. The theory of statistical discrimination suggests that changing the information set of employers at the beginning of the hiring process may have a strong impact on final decisions, *if* skilled minority applicants who are the victims of the negative signal attached to their group are able to overcome that signal at the interview stage by demonstrating their credentials. Similarly, anonymous resumes may be effective against taste discrimination *if* meeting with the applicant induces the recruiter to overcome his prejudiced views against a group, or simply to give these views less weight once they know the individual characteristics of the applicant better. Clearly, little is known on whether these conditions hold in practice. Two non experimental exceptions are Goldin and Rouse (2000) and Aslund and Nordstrom Skans (2007). They both find that anonymous procedures in the first stage of hiring processes increase the relative chances of women to be interviewed and hired.

In the first chapter, we provide experimental evidence on the impact of anonymous resumes on the first and later stages of the hiring process. The experiment was implemented in 2010-2011 by the French Public employment service (PES) to help the French government decide on the enforcement conditions of the law passed in 2006. It was felt that a randomized experiment was needed to provide simple and transparent evidence. Firms posting job offers at the PES were asked to participate to an experiment in which they would have one chance out of two to receive anonymous resumes preselected by the PES agents, rather than standard ones. The experiment involved about 1,000 firms in eight local labor markets, and lasted 10 months. Although the experiment was initially designed to mimic the situation that would prevail if the law was finally enforced, compromises had to be struck given the government's reluctance to impose anything on firms at that stage. In particular, participation to the experiment was not mandatory. This is however the first time, to our knowledge, that experimental evidence is brought on the effectiveness

[3] A worker has a foreign background if he/she is an immigrant or a child of immigrant

of anonymous referral procedures in a large labor market[4], rather than on the mere existence of differential treatment.

The main findings are the following. First, and as expected, women do benefit from higher callback rates under the anonymous resumes procedures; however, the effect is somewhat limited by the extent of labor market segmentation. Indeed, half of the job offers in the experimental sample have only male applicants, or only female applicants. Second, and in a much less expected way, applicants from foreign background or residents in deprived neighborhoods witness a *decrease* in their relative chances to be interviewed, as compared to the reference group. Third, we find evidence that anonymous resumes counter homophily in the hiring process: they undo the tendency of female recruiters to select female applicants, and of male recruiters to select male applicants. Interestingly, this effect persists at later stages of the hiring process, so that anonymous resumes in effect equalize the chances of applicants of both genders to be interviewed and finally hired, irrespective of the gender of the recruiter. Last, we do not find any evidence that the anonymous procedures increases the firms' direct hiring costs nor the opportunity cost of vacancies.

We also document the representativeness of the sample of firms entering the experiment. Although differences in terms of observable characteristics seem minor, there are indications that firms who accepted the experiment were initially rather favorable to applicants from foreign background or residents in deprived neighborhood. This limits the external validity of the experiment. This also provides a plausible interpretation to the counter-intuitive impact of anonymous resumes on that group: these self-selected firms may practice "reverse discrimination" (possibly motivated by many reasons, including the possibility to pay minority workers lower wages), and anonymous resumes may prevent that practice. Interestingly, there is no evidence of selection at entry along lines of gender differential treatment. Less visible in the French debate, gender differential treatment may well be a relevant target for anonymous resumes.

1.1.2 Unemployment insurance

Contrary to anonymous resumes, the economic literature presents a large body of empirical evidence on the impact of unemployment insurance (UI) generosity. Putting aside insurance provision, this literature mostly focuses on impacts on labor market transitions from unemployment to employment. When unemployment benefits are more generous, reservation wages may increase and/or search effort may be lower. This leads to a decrease in unemployment exit rate to jobs. In his seminal work, Meyer (1990) identifies the effect of UI generosity in the US through variations across states. Since the adoption of more generous UI is potentially endogenous at the state level, Card and Levine (2000) propose to focus on exogenous variations in UI generosity due to targeted unanticipated policy change.

[4] Only one small scale experiment has been conducted in the academics market (see Krause, Rinne, and Zimmermann (2011)).

Using the same identifying method, positive effects of potential benefit duration (PBD) on unemployment duration[5] are found in European countries, such as Germany (Hunt (1995)), Austria (Winter-Ebmer (1998), Lalive and Zweimuller (2004), Lalive, Ours, and ZweimÃ¼ller (2006)), Poland (Puhani (2000b)), Slovenia (van Ours and Vodopivec (2006)), Finland (Kyyrä and Ollikainen (2008)) and Portugal (Addison and Portugal (2008)). Other authors rely on discontinuities in the UI system to identify the effects. Those discontinuities are usually age thresholds, as in Lalive (2008) and Caliendo, Tatsiramos, and Uhlendorff (2009). One exception is Card, Chetty, and Weber (2007) who use discontinuities based on past employment thresholds.

At the same time, unemployment benefits may affect the match quality, as it encourages unemployed to wait for higher productivity jobs (see Marimon and Zilibotti (1997) and Acemoglu and Shimer (2000)). Effects on match quality are far less documented (see the review in Addison and Blackburn (2000)). Using a structural model, Belzil (2001) finds that increasing the PBD by one week leads to an increase in subsequent employment duration by 0.5 to 0.8 days. Jurajda (2002) and Tatsiramos (2009) compare benefit recipients to ineligible unemployed and find large positive effects of eligibility on employment duration. Centeno (2004) estimates that a 10% increase of unemployment insurance generosity translates into a 3% increase in subsequent job tenure. In more recent studies, authors focused on identifying causal effects through difference in difference method (van Ours and Vodopivec (2008)) or through regression discontinuities method (Card, Chetty, and Weber (2007), Lalive (2007), Centeno and Novo (2009), Caliendo, Tatsiramos, and Uhlendorff (2009)). They do not find any average effects of PBD on subsequent wage, nor on employment duration. However, Centeno and Novo (2009) and Caliendo, Tatsiramos, and Uhlendorff (2009) document heterogeneity in the effect. Centeno and Novo (2009) find that more constrained unemployed experience an increase by 3 to 8% on their earnings when PDB increases by 6 months. Caliendo, Tatsiramos, and Uhlendorff (2009) find that unemployed persons who find jobs just before their unemployment benefits run out accept less stable jobs than comparable unemployed persons who benefit from longer entitlement.

The second chapter of this thesis provides evidence that effects on match quality are indeed very limited in the French case. Compared to previous studies, this evidence is all the stronger that it concerns workers who are marginally attached to the labor market. Those workers are likely to benefit the most from extended UB. Their marginal attachment shows that they typically lack productive or job search skills that they could acquire with extended UB. They are also likely to be financially constraint such that extended UB would greatly change the value they attach to unemployment.

Our evidence is also all the stronger that we estimate the effect of a large increase in UB generosity. In a regression discontinuity design (RDD) similar to Card

[5]Positive effects of replacement ratios are also found through difference in difference methods in Sweden (Carling, Holmlund, and Vejsiu (2001) or Bennmarker, Carling, and Holmlund (2007)) and in Finland (Uusitalo and Verho (2010)).

Chetty and Weber (see the second section of this introduction for an explanation of the method), we estimate the impact of an increase from 7 to 15 months in potential benefit duration (PBD). In the French unemployment insurance system, when workers work more than 8 months over the year before their job separation, they are entitled to 8 more months of UB: their PBD is more than doubled.

Absence of match quality effect is all the more compelling that extension of PBD actually slows down unemployment exits to work. Unemployed with extended PBD actually wait longer before taking a job (roughly 2.5 months). Yet they do not find better jobs.

Our result is robust to different measures of match quality: employment duration and hourly wage of the first job after unemployment exit. We complement those two standard indicators by the wage two years after unemployment entry. This enables us to compare short and long PBD recipients at the same horizon, whatever the effect of PBD on unemployment duration.

The effect on unemployment exit to employment starts early in the unemployment spell, even when both short and extended PBD unemployed receive benefits. This points to forward looking behaviors. However the effect is somehow stronger between 7 and 15 months after unemployment entry, when short potential benefits are expired but extended benefits are still paid.

1.2 Methodological issues for labor market policy evaluation

Our evaluation exercises adopt the conceptual framework of the treatment effect literature (Rubin (1978)). Treatment evaluation consists in comparing the situation of treated individuals to their situation if they have not been treated. The fundamental problem of evaluation is a missing data problem: this counterfactual situation is not observed at the individual level. Moreover, for most real treatments, eligibility conditions and take up behaviors make the treated and untreated groups fundamentally different, such that untreated individuals are not a suitable control group. LaLonde (1986) documents the extent of this bias in the context of training evaluation. To interpret differences between treated and untreated groups as causal effect of the treatment, randomness must somehow select groups in and out of the treatment. This condition guarantees ex ante statistical comparability between groups.

Randomness can be intentionally generated by researchers to conduct the evaluation exercise (see Duflo, Glennerster, and Kremer (2008)). Intentional randomness, or randomization, is usually thought as the ideal policy evaluation experiment. Indeed we rely on randomized experiments to evaluate the impact of anonymous applications and job search assistance.

However, randomization is not feasible in every context. For example, French law and culture makes it difficult to randomize the generosity of unemployment benefits. To evaluate unemployment benefits extension, we thus rely on local randomness generated by discontinuous eligibility rules. Depending on past employment duration, unemployed are entitled to different potential benefit duration. If they work between 6 and 8 months over the preceding year, they may receive benefits during 7 months; if they work more than 8 months, they are entitled to 15 months of benefits. Thus crossing the 8 month threshold in past employment duration leads to an increase in UB generosity. Potential benefit duration, the treatment, is a discontinuous function of past employment duration, which is called the forcing variable in the Regression Discontinuity Design (RDD) terminology. The key assumption of the RDD evaluation exercise is the absence of manipulation of the forcing variable. In other words, workers are assumed not to sort around the 8 month threshold in past employment duration. This can be a strong assumption. Because they may gain 8 month of UB, workers have strong incentives to work a few days more. Workers reacting to this incentive should accumulate just above the threshold and have thus a higher propensity to receive generous UB. This phenomenon pollutes the comparison between treated and untreated individuals as manipulating workers are also likely to react strongly to UB generosity. Fortunately, some statistical tests can detect manipulation and gauge the plausibility of the randomness assumption, also called exogeneity assumption. In chapter 2, we apply a RDD and do not find evidence of manipulation.

Even when feasible, randomized experiment may suffer from several caveats. In this thesis, we devote specific attention to three of them:

- John Henry effects (or placebo effects)
- differential sample attrition
- equilibrium effects

Placebo effects are discussed while evaluating anonymous applications in the first chapter. The two other caveats are the main problematic of the third and fourth chapter.

Out of those caveats, placebo effects are typical to randomized experiment while differential sample attrition and equilibrium effects are usual criticism of microeconometric evaluation.

1.2.1 Placebo effects

When individuals are aware that they are part of an experiment, whether treated or not, they may have specific experimental behavior that they would not have outside of an experiment. Motivation effect can arise when individuals know that they are observed. For example, because they are in an experiment, individuals exert more effort than usual. Another experimental effect could be that individuals alter their behavior to confirm with their preconceived views of the treatment effect. Suppose that some individuals think that the treatment is ineffective, they may provide less effort when treated than when untreated, thus attenuated treatment effect estimates.

More fundamentally, being in a treatment group is usually perceived positively. This is conceived as an opportunity, whereas being in the control group is a loss of opportunity. Being in the control group is negatively associated. Such assortative association may bias upwards treatment effect.

Experimental effects are usually decomposed in two categories (see Duflo, Glennerster, and Kremer (2008)). When the behavior of the treatment group is altered by the experiment, they are called Hawthrone effects. When it is the behavior of the control group which is altered, they are called John Henry effects. In the anonymous application experiment, John Henry effects are an issue. They are all the more likely that they could explain the counterintuitive effect on foreign vs native interview gap. The control group is composed of employers who have accepted to receive anonymous resumes but receive nominative application. We could suspect that they exert special effort not to discriminate, because they know that their practices are observed and that discrimination is prohibited by the French Law. To test for this placebo effect, the control group has been observed after the end of the experiment. Applicants to post experiment job offers posted by control employers have been surveyed and no difference in interview or hiring gaps have been detected before and after the experiment. To control for confounding factors, employers outside the experiment have also been observed, during the experiment and afterwards. Difference in difference estimates also reject placebo effects.

1.2.2 Differential sample attrition

Sample attrition is a pervasive issue for surveys in social sciences. The damage appears particularly clearly in randomized experiments or quasi-experiments: while randomness in assignment to treatment creates a treatment group and a control group that are comparable, in the presence of sample attrition, however, the observed treatment and control groups may not be comparable anymore, threatening the validity of the evaluation. A variety of tools has been developed to correct for sample selection over the past decades, starting with seminal papers by Heckman (1976 and 1979) and turning less and less parametric up to the "worst-case", assumption-free approach developed by Horowitz and Manski (1995, 1998 and 2000).

The main purpose of the third chapter is to propose another approach to correct sample selection, at the crossroads of semi-parametric forms of the "Heckit" and of the bounding approach of Lee (2009b). The main advantage of our approach is to yield informative set identification without requiring an instrument, but making the most of quite basic information on the number of attempts that were made to obtain response to the survey from each individual that responded. When the number of attempts can be assumed continuous, we show that our approach even yields point identification. Our approach assumes that the distribution of the number of attempts gives information about the ranking of individuals in terms of reluctance to respond and that this ranking is the same whether the individual is treated or not. This assumption boils down to write the response behavior as a latent threshold-crossing model with latent variable independent to the treatment. It can also been restated in terms of monotonicity and compared to Lee's assumption. Our assumption being stronger and our information set been larger, our bounds are tighter than Lee's bound.

We then apply our sample selection correction in the context of the Behaghel, Crepon, and Gurgand (2012) job search experiment, which can be viewed as a within-study comparison proving that sample selection can matter in practice. In the context of that job search experiment, exhaustive administrative is available, but a phone survey yields richer information – with low and unbalanced response rates. Using the administrative information, we show that selection into the phone survey is not as good as random, as it is correlated to potential outcomes. Moreover, point estimates suggest that the phone survey over-estimates the program's impact by about 50%. Applying the sample selection correction procedure closes most of the gap between the estimates in the full and in the selected samples. Bounds *à la* Horowitz and Manski (2000) or Lee (2009) are, in this application, too wide to be very conclusive.

1.2.3 Equilibrium effects

External validity of empirical results may be compromised in some randomized experiment or any microeconometric evaluation because of equilibrium effects. Even if treated and untreated individuals are ex ante identical, the presence of treated individuals may affect the untreated individuals, contradicting the so-called 'nointerference' (Rubin (1978)) or 'stable unit treatment value' (Angrist, Imbens, and Rubin (1996)) assumption. Then difference in outcomes between treated and untreated measures the effect on the treated net of that on the untreated. When the proportion of treated individuals changes, the difference in outcomes is also likely to change. As a consequence, tone particular microeconometric estimate is unlikely to be valid when the treatment is implemented in another environment at a different scale. To obtain valid estimate, one solution is then to change the scale of randomization such that randomized units do not interfere with one another. This is usually a hard task when evaluating labor market policies, as this solution means finding numerous segmented labor markets. Another approach is to try to estimate the extent of the equilibrium effects as in Crepon, Duflo, Gurgand, Rathelot, and Zamora (2012) by varying the proportion of treated individuals in different labor markets.

The issue of equilibrium effects, which is discussed in a broader perspective in the survey of Meghir (2006), is particularly relevant to the evaluation of labor supply based policies (such as increasing incentives or monitoring the unemployed). First, they generally aim at increasing the overall number of filled jobs, which depends on the interactions between aggregate labor supply and labor demand. Second, these policies may induce displacement effects: treated persons may crowd out the untreated because they compete for the same jobs.

Although they have long been recognized, these questions have received limited attention to date. Davidson and Woodbury (1993) and Calmfors (1994) are early contributions. More recently, Lise, Seitz, and Smith (2005) study the equilibrium effects of the Self-Sufficient Project incentive program in Canada. They calibrate an equilibrium model of the labor market so that, when used in partial equilibrium, the model matches the effect of the program estimated by direct comparison of treated and untreated. When equilibrium effects are simulated, the impact of the Self-Sufficient Project is far lower. In contrast, Albrecht, van den Berg, and Vroman (2009) find, using a calibrated model, equilibrium effects of a Swedish training program to be stronger than implied by direct comparison. Using a job search and matching model with skilled and unskilled workers, Van der Linden (2005) shows that micro and equilibrium evaluations are likely to differ widely when job search effort and wages are endogenous. When wages are bargained over, raising the effectiveness of or the access to counseling programs pushes wages upwards and leads to lower search effort among nonparticipants. Induced effects can outweigh positive micro effects on low-skilled employment when the response of wages is taken into account.

The equilibrium effects have also been analyzed in empirical evaluations that do not rely on structural models. For instance, the contribution of Blundell, Costa Dias, Meghir, and J. (2004) evaluates the New Deal for Young People in the U.K. This program was piloted in certain areas before it was rolled out nation wide. Moreover, the program has age specific eligibility rules. Blundell, Costa Dias, Meghir and Van Reenen use these area and age based eligibility criteria that vary across individuals of identical unemployment durations to identify the program effects. They find that either equilibrium wage and displacement effects are not very strong or they broadly cancel each other out.

The aim of the fourth chapter is to analyze the impact of counseling in the standard matching model of the labor market (Pissarides, 2000). In our specification, counseled unemployed have a constant comparative advantage in the job search. This is consistent with most of empirical results on Job Search Assistance - JSA - evaluation (see meta analysis in Kluve (2010) and Card, Kluve, and Weber (2010)). According to Kluve (2010) , out of 21 microeconometric evaluations of JSA in European countries, 15 conclude on short term positive significant effects on transitions out of unemployment and 6 on insignificant effects.[6] Using this simple matching model allows us to analyze the consequences of counseling in a dynamic set-up. More precisely, we shed some light on three important issues:

(i) What is the true impact of the policy when equilibrium effects are taken into account? The model shows that the true impact of counseling can be very different from what can be concluded when equilibrium effects are neglected even when the

[6] In the French case, Fougère, Pradel, and Roger (2009) find that standard counseling by the Public Employment Service increases unemployment exit rate. Crépon, Dejemeppe, and Gurgand (2005) find that intensive counseling increases by 20% the unemployment exit rate with respect to standard service by the PES. Based on experimental evidence, Behaghel, Crepon, and Gurgand (2012) also find that intensive counseling increases employment rate after 8 months since unemployment entry by 4 to 9 points.

treatment group is small. For instance, we find that counseling can increase unemployment when a small proportion of job seekers benefit from counseling, although counseling improves the efficiency of job search. Equilibrium effects rely on the adjustment of wages. The impact of policies on wages has been analyzed in some papers devoted to equilibrium effects of several labor market policies and education policies, in particular since the seminal contribution of Heckman, Lochner, and Taber (1998). Our model allows us to analyze precisely the reaction of wages to counseling, as in the paper of Van der Linden (2005).

(ii) What is the impact of the generalization of the policy to a large treatment group? The model shows that there is no simple answer. In particular, the relation between the impact of the policy on unemployment and the size of the treatment group is not necessarily monotonic. Strikingly, in our framework, unemployment increases with the size of the treatment group when a small share of job seekers are treated but diminishes with the size of the treatment group when a sufficiently large share of job seekers are counseled.

(iii) What is the dynamic impact of counseling? Many experiments made to evaluate labor market policies are transitory. Typically, a group of job seekers is selected to benefit from counseling (the treatment group) and the control group will never benefit from counseling. The comparison between the outcomes yields the evaluation of the impact of counseling. Our model allows us to stress that the consequences of permanent and transitory policies can be very different. The difference comes from the reaction of non-counseled job seekers. When the policy is transitory, non-counseled workers do not expect to benefit from counseling in the future. However, when the policy is permanent, the expectation to benefit from counseling in the future induces the non-counseled workers to raise their reservation wage. In our framework, this phenomenon implies that permanent counseling increases unemployment when a small share of job seekers are counseled whereas counseling always decreases unemployment when it is transitory. Accordingly, it can be misleading to conclude that a truly successful transitory policy will remain successful when it becomes permanent.

Chapter 2

Do anonymous resumes make the battlefield more even? Evidence from a randomized field experiment[1]

Contents

2.1 Introduction

Differential treatment by race and gender are repeatedly documented as a prominent feature in many labor markets, despite decades of anti-discrimination laws

[1] This chapter is largely based on common work with Luc Behaghel and Bruno Crepon.

that explicitly prohibit and strongly penalize such firm behavior.[2] This persistence has sparked the debate on alternative (or complementary), non coercitive policies against discrimination. Anonymous referral procedures have received lots of attention. Anonymous resumes are used in Belgium in the federal administration. They have been tested locally in the Netherlands, in Sweden and in Switzerland, but the tests have not led to new legislation. In Great Britain, an attempt by Liberal Democrats to impose anonymous resumes for recruitment has been opposed by the government in 2009. In France, which we study here, a law was passed in 2006 that made the use of anonymous resumes mandatory for firms with more than 50 employees; yet, the government did not take the steps to define the conditions under which the law would apply.

In the absence of strong empirical evidence, the confusion and hot debates around anonymous resumes are understandable: there are strong, plausible pros and cons. The theory of statistical discrimination suggests that changing the information set of employers at the beginning of the hiring process may have a strong impact on final decisions, *if* skilled minority applicants who are the victims of the negative signal attached to their group are able to overcome that signal at the interview stage by demonstrating their credentials. Similarly, anonymous resumes may be effective against taste discrimination *if* meeting with the applicant induces the recruiter to overcome his prejudiced views against a group, or simply to give them less weight once they know the individual characteristics of the applicant better. Clearly, little is known on whether these conditions hold in practice. While proponents and opponents of anonymous resumes usually agree that the measure should change the pool of candidates called for an interview, they hold opposite views on whether this change would be sufficient to overcome discrimination in later stages of the hiring process, once the identity of the applicant is revealed to the employer. Additional arguments in the debate concern the cost of the measure: by removing information, the measure can be viewed as increasing matching frictions on the labor market, with ultimately negative welfare impacts on firms and workers.

This paper provides experimental evidence on the impact of anonymous resumes. The experiment was implemented in 2010-2011 by the French Public Employment Service (PES) to help the French government decide on the enforcement conditions of the law passed in 2006. It was felt that a randomized experiment was needed to provide simple and transparent evidence. Firms posting job offers at the PES were asked to participate to an experiment in which they would have one chance out of two to receive anonymous resumes preselected by the PES agents, rather than standard ones. The experiment involved about 1,000 firms in eight local labor markets, and lasted 10 months. Although the experiment was initially designed to

[2] Correspondence testing studies, initially primarily developed in the UK, have been instrumental in providing direct and compelling evidence, even though the exact interpretation in terms of discrimination is still debated. A notable example for race in the US is Bertrand and Mullainathan (2004). Correspondence studies have been increasingly used in France, the country under study here: see in particular Duguet, L´ Horty, Meurs, and Petit (2010) introducing the special issue of *Annals of Economics and Statistics*, n°99/100, on measuring discrimination.

mimic the situation that would prevail if the law was finally enforced, compromises had to be struck given the government's reluctance to impose anything on firms at that stage. In particular, participation to the experiment was not mandatory. This is however the first time, to our knowledge, that experimental evidence is brought on the effectiveness of anonymous referral procedures, rather than on the mere existence of discrimination.

The main findings are the following. First, and as expected, women do benefit from higher callback rates under the anonymous resumes procedures; however, the effect is somewhat limited by the extent of labor market segmentation, as half of the job offers in the experimental sample have male applicants, or female applicants only. Second, and in a much less expected way, applicants from foreign background or residents in deprived neighborhoods witness a *decrease* in their relative chances to be interviewed, as compared to the reference group. Third, we find evidence that anonymous resumes counter homophily in the hiring process: they undo the tendency of female recruiters to select female applicants, and of male recruiters to select male applicants. Interestingly, this effect persists at later stages of the hiring process, so that anonymous resumes in effect equalize the chances of applicants of both genders to be interviewed and finally hired, irrespective of the gender of the recruiter. Last, we do not find any evidence that the anonymous procedures increases the firms' direct hiring costs nor the opportunity cost of vacancies.

We also document the representativeness of the sample of firms entering the experiment. Although differences in terms of observable characteristics seem minor, there are indications that firms who accepted the experiment were initially rather favorable to applicants from foreign background or residents in deprived neighborhood. This provides a plausible interpretation to the counter-intuitive impact of anonymous resumes on that group: these self-selected firms may practice "reverse discrimination" (possibly motivated by many reasons, including the possibility to pay minority workers lower wages), and anonymous resumes may prevent that practice. Interestingly, there is no evidence of selection at entry along lines of gender discrimination. Less visible in the French debate, gender discrimination may well be a relevant target for anonymous resumes.

The next section relates our evaluation to the relevant literature. The following sections present some theoretical insights, the experiment, the data, and the measures used to characterize groups at risk of discrimination. A specific section is devoted to analyzing the representativeness of firms participating to the experiments. The last two sections present the results from the perspective of applicants and firms, respectively.

2.2 Previous literature

Despite the well-documented widespread discrimination in hiring and the political will to fight against it, there are relatively few evaluations of anonymization during the recruitment process. To our knowledge, two notable exceptions are Goldin and

Rouse (2000) and Aslund and Nordstrom Skans (2007). In both evaluations, the introduction of anonymization during the recruitment process is found to increase the hiring rates of women relative to men.

Goldin and Rouse (2000) analyze the introduction of shields in hiring auditions of American Philharmonic orchestras. They identify the effect of shield adoption in a difference and difference framework, assuming that shield adoption is not simultaneous to any changes in other anti-discriminatory practices in the orchestras. Thanks to randomization our evaluation does not rely on such assumptions. They find that women have a higher probability to advance to later stages of the recruitment process when shields are used. Moreover, even if later stages are not anonymous any more, women have a higher probability to be hired if first stages are blind. One possible interpretation of their findings is that, knowing that auditions are blind, more talented women applied to the job opening. In other words, adopting a blind recruitment process sends signals to potential candidates who self select out of the market when the process is nominative. Blind auditions not only change the information set of recruiters in early stage of the recruitment process, but it could also change the composition of the pool of candidates applying. Our experimental design mitigates this 'calling' effect and enables us to estimate the pure information effect.

Aslund and Nordstrom Skans (2007) evaluate the effects of anonymous application forms introduced in the recruitment of 109 public jobs. Those jobs were advertised as anonymous and candidates had to follow a specific application procedure. The evaluation may also estimate the 'calling' effect. The anonymous applications were experimented in two voluntary districts of Goteborg city in 2004-2006. As Goldin and Rouse, they use a difference in difference framework and find that the probability of being interviewed and hired are equalized between male and female candidates when applications are anonymous. They also find that the interview rate is leveled between candidates with foreign origin and natives, but the hiring rate of natives is still higher under anonymous applications. This is first evidence that the efficiency of anonymous procedure is heterogenous, evidence which we also confirm.

Among the wide literature on discrimination, our article contributes to another empirical strand which focuses on what can be called "homophily" or own-group bias. Behind this concept is the simple idea that human beings tend to prefer to interact with people from the same ethnic group, the same gender... This behavior can reveal a true preference, in this case homophily can be associated to taste-based discrimination, or it can be simply rational : obtaining relevant information - extracting a signal - is easier from someone of the same ethnic/gender group (statistical discrimination). Price and Wolfers (2010) find that more personal fouls are awarded against players when they are officiated by an opposite-race officiating crew than when officiated by an own-race refereeing crew. Anwar and Fang (2006) find that troopers from different races are not monolithic in their search behavior. However the authors do not reject the hypothesis that troopers of different races do not exhibit relative racial prejudice. Fisman, Iyengar, Kamenica, and Simonson

(2008) document same-race bias in the marriage market (Speed Dating experiment) and their determinants, such as gender, social background, age... We contribute to this empirical literature by documenting such bias on the labor market and by extending the usual own-ethnic bias analysis to own-gender bias.

2.3 Theoretical insights

We present a parsimonious model to highlight expected effects of anonymous resumes. This recruitment model has two steps, interview and hiring, with incomplete information at the first stage. To make our description more convincing, we choose a fully rational model, recruiters do not change their a priori when meeting candidates[3]. A model with such inconsistent behavior could have any predictions about effects of anonymization.

2.3.1 Environment

There are 2 types of candidates : from the majority group (0) and from the minority group (1). Each type has its own distribution of net profit (p) to the firm : $F_0(p)$ and $F_1(p)$; p can represent:

- productivity if wages are fixed and equally distributed among groups;
- net output if wages are different (typically lower for the discriminated group)
- profit net of disutility; the recruiter may have different taste to work with one or the other type

The recruiter has a prior on π the proportion of candidates from the minority. The recruitment is in 2 steps:

1. The recruiter receives N resumes. He does not observed p. He chooses M candidates to interview.
2. At the interview stage, which is costly (c per candidate), the recruiter observes p. He chooses randomly to recruit any candidate whose productivity is above $\bar{p}$

2.3.2 When resumes are nominative

The recruiter knows the resumes' types. He receives N_0 and N_1 resumes. He does not observed p. He chooses M_0 and M_1 candidates of each type to interview. The probability that all candidates would yield a profit less than $\bar{p}$ is $(F_0(\bar{p}))^{M_0}\,(F_1(\bar{p}))^{M_1}$. To choose among the N nominative resumes, the recruiter maximizes:

$$\bar{p}\left(1-(F_0(\bar{p}))^{M_0}\,(F_1(\bar{p}))^{M_1}\right)-c(M_0+M_1)$$

[3]With the following notations, in such a model, the distribution $F(.)$ used to choose interviewed candidate would be different than the one to choose hired candidates.

such that $M_0 \leq N_0$ and $M_1 \leq N_1$. Suppose $F_0(\bar{p}) < F_1(\bar{p})$, the recruiter chooses first the resumes from the majority group and, if there is still some net gain to expect, chooses among the minority group.

Suppose all recruiters are similar and have all N_0 and N_1, two types of aggregate states can arise: one with strong discrimination (members from the minority are never interviewed), one with mild discrimination... In reality, among the N resumes, some recruiters receive few minority candidates (low N_1) and some many of them (high N_1). When there are only few minority candidate in the initial pool, they are never called for an interview (even if they are highly productive) : discrimination is strong. When there are only a few majority candidate, some minority candidates are called for an interview, discrimination is mild.

2.3.3 When resumes are made anonymous

The recruiter faces now a distribution $F(p)$. We assume that $F_0(\bar{p}) < F(\bar{p}) < F_1(\bar{p})$. He chooses candidates indifferently of the type and the probability to be interviewed is the same in both groups. Some minority candidates who would not have access to the interview stage when resumes are nominative gain access in the treatment group.

The number of candidates interviewed in the treatment group can be higher or lower than in the control group depending on the composition of the initial pool. When there are many majority candidates, the number of interviews would be lower in the treatment group.

Even if the number of interviews is the same in both treated and control groups, the distribution of the composition of interviewed pool is affected by the treatment. Minority candidates are more often in competition with other minority candidates. This alters the relative chance to get hired for a minority candidate.

This simple model predicts that chances to be interviewed are equalized between majority and minority candidates when resumes are anonymous. But the effect on the average interview rate can be ambiguous. In this model, conditional on the composition of the pool of interviewed candidates, the relative chances of minority and majority candidates to get hired are not altered by anonymous resumes. The differential treatment at this later stage is not affected. However because the pool of interviewed candidates is stronger in minority candidates when resumes are anonymous, the gap of the unconditional hiring probability is reduced. The extent of this reduction depends of course on the difference between F_0 and F_1. If F_1 is strongly dominated, actual effects are likely to be small and economically insignificant.

Note that, when resumes are nominative, our model exacerbates differential treatment, whatever their nature. On top of differential treatment prevalent at the hiring stage, some kind of statistical discrimination is induced by information incompleteness at the first stage. When differential treatment is taste based (p is productivity net of the recruiter's disutility), statistical discrimination enhances taste based discrimination.

2.4 Experimental design

In this section, we present the experimental design used to measure the impact of anonymous resumes. The experiment was conducted in 8 (out of one hundred) French *départements*, at branches of the public employment service (PES) located in urban areas, during 10 months. It proceeded as follows:

1. **Firm entry in the experiment.** Firms posting job offers at the PES have the option to ask for a PES agent to make a first screening of applicants based on their resume. In that case, the firm receives only selected resumes from the PES (from a couple to a dozen, in most cases), instead of having applicants contact them directly. This service is free. During the time of the experiment, all firms with more than 50 employees posting a job lasting at least 3 months and asking for this service are invited to enter the experiment. They are told that their job offer will be randomly assigned to the anonymous or standard procedure, with probability 1/2. Firms are free to refuse; however, in order to induce positive responses, participation is presented as the default option. A given plant enters the experiment at most once: plants that have already entered the experiment are no longer asked to participate.

2. **Matching of resumes with job offers.** The job offer is posted by the PES on a variety of supports, including a public website on which interested job seekers are asked to apply through the PES branch. The PES agent selects resumes from these applicants and from internal databases of job seekers. A first lot of resumes is thus matched with the job offer.

3. **Randomization and anonymization.** The resumes are sent to research assistants in charge of the randomization at the central PES offices. Job offers (and their first lot of resumes) are randomly assigned (using a random number generator) to treatment or to control group, with probability 1/2. If the offer is assigned to the treatment group, all the resumes are given a number and anonymized by the research assistant[4]; then, they are sent back to the PES agent in charge of the job offer follow-up.

4. **Selection of resumes by the employer.** The employer selects the resumes of applicants she[5] would like to interview. Control group employers contact the applicants directly, treatment group employers give the PES agent the resumes' numbers so that it is the PES agent who sets up the hiring interviews, in order to maintain the applicant anonymity.

5. **Additional lots of resumes.** If the employer could not fill the position with the first lot of resumes, she requests additional lots. The PES sends a new lot of selected resumes with the same format as for the first lot.

[4]The degree of anonymization is described below.

[5]As shown below, most hiring officers in the experiment are females; we will therefore use feminine pronouns.

This experimental design calls for a few comments:

Plants enter the experiment at most once, either in the treatment or in the control group. The main reason was the fear of the PES that repeated participation to the experiment and the corresponding surveys would have been too much trouble for firms. To maximize positive responses when inviting the firm to participate in the experiment, it was therefore clearly specified that the experiment would only concern one job opening. It could also be argued that having the same hiring officer acting in turn as treatment and as a control individual would have made the results harder to interpret, as this would have increased the risk that her behavior be affected by her previous participation. Possible Hawthorne or Henry effects are discussed in the results section. A drawback of this is that the experiment does not capture learning effects nor the long-term impact of using anonymous resumes.

Anonymization is limited. Anonymization consisted in erasing the top part of the resume: name, address, gender, ID picture, age, marital status and number of children. However, it did not imply any further standardization of the content of the resume. In particular, information on gender could be read from gender-specific terms used in the main part of the CV; neighborhood of residence could be partly inferred from information on where the applicant graduated from high-school; and ethnicity could be spotted from foreign language skills. Going further would have implied much more complex logistics during the experiment, and it was felt that standardization would anyway not have been feasible if anonymous resumes had been made mandatory nationwide.

Randomization occurs at the job offer level. For a given job offer, all resumes transiting by the PES are treated identically (either anonymous, or standard). This level of randomization corresponds to the policy evaluated, that would have all resumes anonymized, instead of some anonymous resumes competing with standard resumes. However, the PES is not the only channel for recruitment: firms may also receive applicants from other sources, whose resumes are not anonymized. We measure below whether firms substitute these other channels to the PES in response to anonymization.

Randomization occurs after matching resumes to job offers. Had the randomization occurred after randomization, the PES agent could have selected different applicants for job offers with anonymous resumes (consciously or not). This would have affected the comparability of treatment and control applicants. To avoid this, a first lot of resume was selected before randomization occurred. Most analyses below are restricted to these first lots, as they contain resumes that are by construction statistically identical in the control and treatment group. We check below whether resumes in subsequent lots sent by the PES agent differ from the first lots.

To summarize, the goal of this experimental design is to mimic as closely as possible what making anonymous resumes mandatory would change for recruitment. By contrast with a law that would have anonymous resumes mandatory, there are however two main caveats: first, only a fraction of the targeted job offers entered the experiment, as the experiment was run in specific urban areas and employers were allowed to opt out; second, only applicants transiting by the PES were concerned, as the firms could keep using their other (non anonymous) recruitment channels. This results from the constraints set by policy makers when launching the experiment, despite our attempts to make participation to the experiment compulsory for all firms using the PES. As detailed below, the data collection strategy was adapted to measure the consequences of these features of the design.

During the ten months of the experiment (November 2009 to September 2010), 1,005 job offers entered the experiment out of total of a bit more than 6,000 eligible offers (each plant counting for one offer). This limited entry into the experiment is due to losses at two steps. First, using administrative data on all job offers posted at the PES, one can check that only 25.5% of the eligible employers were invited to enter the experiment. It should be noted that the experiment took place at a time when inflows of job seekers were extremely large due to the recession, so that PES agents were extremely busy and some of them simply forgot or neglected to invite firms to participate. However, it is also likely that some PES agents preferred not to invite firms that they expected would refuse. Among firms invited to enter the experiment, the take-up rate amounts to 63.3%. Clearly, although only 37.7% of firms formally declined to participate, the representativeness of the experimental sample is an issue, and it is analyzed in depth in section 2.7.

2.5 Data

We collect administrative and survey data.[6] The administrative data covers all firms and all job seekers who used the public employment services in the experimental areas during the experiment. It has basic information on the firm (size, sector), the job position offered (occupation level, type of contract), limited information on the job seeker (unless the job seeker has a file as unemployed). It also provides a follow-up of the recruitment process until the position is filled or the job offer is withdrawn; however, the quality of that follow-up is weak, and some critical information is missing (in particular, one does not know whether the candidate was interviewed before the firm rejected his application). In what follows, the administrative data is mostly used to characterize the population of firms entering the experiment, by comparison with the broader population of firms interacting with the PES.

We conducted telephone interviews with all firms entering the experiment, as well as with a subsample of applicants to these firms. The data from these two surveys constitute the core database used in the analysis. In addition, we interviewed a sample of firms that had refused to enter the experiment or that had not been invited by PES agents, despite the fact they were eligible for the experiment: again, the goal is to check whether our core sample is representative of the target population of firms. Last, a subsample of applicants on job offers from control group firms *after the experiment* was also interviewed: as detailed below, the goal is to check whether control firms behaved in a specific way during the experiment. The

[6] In addition to these two main sources, information available in the resumes was also coded.

surveys used for applicants (respectively, for firms) were similar across subsamples. We now present these two surveys briefly; specific questions will be presented when they are used in the analysis, and survey tools (in French) are available on line.

The main goal of the survey of applicants is to provide a reliable measure of whether the applicant was interviewed for the job, and of all his characteristics that could lead to discrimination. We ask in particular for the country of birth and the citizenship at birth, both for the applicant and his parents. There are also questions about the applicant's labor market situation, the recruitment process, as well as subjective questions on self-confidence, perceived discrimination, and perceived labor market prospects.

The firm survey has three main functions. The first one is to measure the result of the recruitment, in particular when the recruitment was abandoned without filling the position, or when the hired candidate came from other channel than the PES (in which case he would not be present in the survey of applicants). Second, the survey includes detailed questions on the hiring process: what were the different steps, how formalized were they, how much time was spent on each of them, who was involved within the firm. Last, the survey tries to characterize the background of the hiring officer who led the hiring process. Just like the applicant, we ask for her country of birth and citizenship at birth, as well as her parents'.[7] In addition, to characterize her social networks and the firm's social composition, we ask for the first names of five friends outside the firm and five colleagues within the firm.

Table 2.1 details the sample of applicants. The initial population (6742 applicants) is partitioned in two ways: control vs. test; at risk of discrimination vs. other. At that stage, applicants at risk of discrimination are identified from the administrative information as those living in a deprived neighborhood or with an African or Muslim-sounding name. They are given higher sampling weights, in order to maximize statistical power. Overall, response rates are around 65-70%; even though they are lower in the control group, the difference is not statistically significant (the p-value is .27). The survey thus yields a total sample of 1977 applicants. Among those, 1260 belong to the first lot of resumes matched to a job offer before the randomization took place. As discussed in section 2.4, these 1260 applicants constitute the cleanest comparison groups; unless otherwise specified, they constitute the sample of analysis.

Table 2.2 presents the sample of firms. There are five separate groups of firms. 385 control and 366 treatment firms accepted the experiment and went through the randomization. 254 firms accepted the experiment but were not randomly assigned to treatment or control: they canceled or filled the job opening before a first lot of resumes was collected and randomization could take place. This underscores the fact that many firms actually fill their positions quickly without any help of the PES. 608 firms refused the experiment, and 4714 were not invited to participate. These last two groups of firms were sampled with lower sampling rates. Their response

[7] Special care was devoted to survey the person in charge of the recruitment. All respondents to the firm survey reported being in charge of the selection of resumes; 89% took part to job interviews.

rates are also somewhat lower, as could be expected. The response rate difference between control and treatment firms is not statistically significant.

2.6 Measuring applicants' risk of discrimination

The purpose of anonymous resumes is to protect potential victims of discrimination by hiding characteristics that would allow firms to screen them before the interview. Discrimination can however occur along many dimensions: ethnicity and foreign origin, neighborhood of residence, gender, age. This section details how we measure these different dimensions.

Gender, age and neighborhood of residence are available in the administrative data; they are also directly reported on the resume. One issue with age is that it can be inferred fairly easily from the content of the resume (in particular, the year the applicant entered the labor market or finished her education): on the basis of this information, it is possible to predict the applicant's age within a four-year bracket in 60% of applications. Removing the exact age could therefore only matter in so far as employers attach a particular signficance to some age thresholds, for instance, the age of 50. We use the corresponding indicator variable in the analysis, but find little impact on the effect of anonymous resumes. We will therefore not focus on age in the analyses that follow.

Another issue is how to characterize deprived neighborhoods of residence. In the US, Bertrand and Mullainathan (2004) use a variety of criteria based on the fraction of Whites, the fraction of college graduates or the average per capita income. We use administrative classifications of neighborhood defined to target subsidies or tax exemptions[8]: their boundaries closely match socioeconomic geographical disparities; moreover, one of their alleged perverse effect is to create a stigma effect. They are therefore particularly relevant to assess the impact of anonymization.

The main issue is how to measure discrimination risk associated with foreign origin or ethnicity. French law forbids the use of ethnic categories that would label someone as White, Black, or African-French, for instance. Instead, we follow a twofold approach. First, in the spirit of correspondence testing studies (see Bertrand and Mullainathan, 2004), we code whether the applicants' names has a foreign-sounding origin. Following research by Felouzis (2003) and Ores (2007), we use the etymology of the applicant's name: Muslim first names are identified from a database created by Chebira (2005). The second approach uses the place of birth and the citizenship at birth. Immigrants are defined as those born abroad who did not have French citizenship at birth. Children of immigrants are those whose father was born abroad and did not have French citizenship at birth. Specific questions are used for the special case of individuals from former French colonies, who might declare they were French citizens at birth if they were born before independence; they are classified as foreigners if they took the citizenship of their new country at

[8]They are known as "Zones urbaines sensibles" (ZUS) and "quartiers en contrat urbain de cohésion sociale (CUCS)"; these zoning schemes are comparable to "Enterprise zones" in the US.

independence. The two approaches – based upon the first name or the migration status – are complementary. In some cases, a foreign-sounding name is the only signal that appears on the resume. But in other cases, immigrants may have a French-sounding name although their origin can be inferred from other signals on the resume (for instance, their last name or an ID picture).[9]

Table 2.3 compares the different measures of discrimination risk. The sample is balanced between men and women; it is clearly skewed toward young candidates. Roughly one applicant out four lives in a deprived neighborhood; the same proportion has a Muslim or African-sounding name; one out of five is immigrant, and that proportion goes to four out of ten for immigrants or children of immigrants. The different measures of origin are correlated. Of particular interest is the correlation between the name and the migration status, shown in table 2.4: clearly, African or Muslim-sounding names correspond to applicants with a foreign origin; however, a significant fraction of immigrants (including those from Africa) do not have an African or Muslim-sounding name. The variables based on immigration (as declared during the interviews) may better capture the risk of discrimination, when that origin can be inferred from other signals in the resume. In the analysis, we compare the effects of using these alternative measures.

Table 2.6 shows no significant observable differences between control and treatment applicants, in the first lot of resumes (selected by the PES before randomization).[10]

Last, table 2.5 displays the average credentials of the different groups of applicants. Specifically, each line corresponds to the regression of a given characteristics (e.g. years of education) on four indicator variables characterizing the applicant's gender, neighborhood of residence, and migration status (distinguishing immigrants and children of immigrants).[11] Applicants from potentially discriminated groups do differ by some observables from the reference candidate (a male who is not an immigrant nor the son of an immigrant, and who does not reside in a deprived neighborhood). Overall, people at risk of discrimination in the sample are younger, have less work experience (in particular in the type of job they are applying for), and tend to have a lower reservation wage. Women and immigrants are more edu-

[9] Alternative measures of origin include the applicant's patronyme and his mother tongue. Measures using the applicant's patronyme were hard to implement and did not seem, by cursory look at the resumes, to improve on the information yielded by the first name and the migration status. Moreover, in the French context, the mother tongue does not allow to capture immigrants well: according to Simon (1998), only 13% of 2nd-generation Algerian youth declare their parent's language as their mother tongue.

[10] We also tested whether differential selection by the PES agent introduces systematic differences between applicants in the treatment and control groups for lots of resumes that were selected after randomization (as would be the case if the agent decided to over-select applicants at risk of discrimination for the anonymous procedure, for instance). There is however no evidence of this: control and treatment applicants remain comparable. More precisely, one does indeed note that resumes from the first lot differ from resumes of the subsequent lots, but the difference is the same for treatment and control job offers. (Results omitted here.)

[11] This additive specification turns out to be a convenient summary. Other descriptive approaches lead to similar main facts.

cated, whereas children of immigrants and residents in deprived neighborhood are less educated; driving licenses are less frequents except for residents of deprived neighborhoods. Overall, this does not suggest that applicants from groups at risk of discrimination have significantly lower credentials. This feature of our sample may of course result from the screening of applicants by the PES.

2.7 Representativeness of firms entering the experiment

Before analyzing the impact of anonymous resumes, it is important to check whether firms entering the experiment are representative of firms targeted by the law. Indeed, as noted above, the experimental design allowed firms to refuse to participate, and a significant fraction (around 38%) did so; moreover, a large share of firms eligible for the experiment were not invited to participate. An obvious question is therefore whether firms that entered the experiment were more or less prone to discriminate than other firms. Different hypotheses are possible. One may suspect that firms that do not discriminate are more likely to accept the experiment. In that case, the evaluation would yield the impact of anonymous resumes on "well-behaved" firms, and would not say anything of their impact on firms that do discriminate. But the opposite may be true. There is anecdotal evidence of firms with a strong policy against discrimination that refused to participate, claiming that anonymous resumes are a heavy procedure and unnecessary procedure, possibly even counter-productive by preventing the firm to take into account the disadvantaged background of applicants when assessing their credentials. Moreover, firms that discriminate may choose to participate to the experiment, by fear of raising suspicions if they did not participate. In that case, the evaluation would estimate the local effect on anonymous resumes on more discriminatory firms (possibly overestimating the average impact on the overall firm population).

To address this question, we take a twofold approach. First, we look for observable differences between firms in and outside of the experiment. Table 2.7 shows that firms participating to the experiment indeed display some specific features, although the differences are not massive. The first two columns describe firms in the control and treatment groups, respectively. As expected with random assignment, differences are small and only one is statistically significant (significant differences are signalled in columns 6 to 9). The third column describes firms that withdrew their offer before randomization could take place. Column 4 (respectively, 5) displays firms that refused to participate (respectively, were not invited to participate). The size and industry of firms that refused to participate are close to those of control firms. But firms that were less frequently invited to participate are concentrated in the non-merchant service sector.[12] Firms refusing to enter the experiment are

[12]One likely explanation for that is that subsidized jobs were excluded from the experiment while these jobs are more frequent in the non-merchant service sector; even though we exclude the corresponding job offers from the table, we were told that some PES agents misunderstood the

less frequently firms offering skilled jobs. Similarly, firms that are not invited to participate are less often offering indefinite duration contracts.

Tables 2.8 and 2.9 complement table 2.7 using the richer information provided by the firm survey (at the cost of reduced sample size, which reduces the likelihood of detecting statistically significant differences). Firms refusing the experiment or not invited to participate less often declare to have mobilization actions against discriminations. Firms that refuse also more often declare having difficulties to fill a vacancy: this may be one reason for not participating, by fear of jeopardizing a difficult recruitment process. More surprisingly, firms that refuse the experiment are also more often frequent users of the public employment services.

All these differences are suggestive of selective entry in the experiment. However, there is no evidence that this selection is correlated with discriminatory practices. In particular, taste-based models of discrimination emphasize prejudiced "tastes" of customers, coworkers, or employers. However, the fact that the position offered implies frequent customer contact or teamwork with coworkers does not correlate with the firm's decision to enter the experiment. Observing employer's tastes is hard, but one can use the detailed information on the hiring officer. Her origin and migration status, her professional or personal networks do not correlate with the entry of the firm in the experiment. In particular, there is no evidence that the composition of the firm or the personal network of the hiring officer – as measured by the presence of African or Muslim sounding names – are different in firms that do not participate to the experiment. Overall, there is evidence that firms entering the experiment are specific, but it is hard to say whether these specificities are linked to discriminatory practices.

The second approach to assess selective entry in the experiment is to look directly at the firms' record in selecting applicants. Unfortunately, interview rates of different groups of applicants are not well measured in administrative sources. Our approach is therefore to extend the survey of applicants – initially designed for applicants to experimental job offers – to a subsample of applicants on job offers that did not enter the experiment (either because the firm was not invited to participate, or because it refused to participate). This allows us to measure interview rates across different types of applicants, and to compare these differences across firms inside and outside the experiment, when using standard (nominative) resumes. The goal is to check whether minority candidates tend to be in a better or worse relative position in firms that entered the experiment. As discussed further in the next section, a parsimonious model to answer that question is:

$$Y_{ij} = a_0 + a_1 M_i + a_2 F_i + d_0 P_j + d_1 P_j \times M_i + d_2 P_j \times F_i + e_{ij}, \qquad (2.1)$$

where Y_{ij} is an indicator variable equal to 1 if applicant i on job offer j is interviewed (or is hired), M is an indicator for being in the group of immigrants, children of immigrants and/or residents of deprived neighborhoods, F is an indicator for female applicants, and P is the indicator variable for the firms participation in the

rule and did not propose the experiment to any firm from that sector, even when the job was not subsidized.

experiment. Testing $d_1 = d_2 = 0$ amounts to testing whether the relative chances of potentially discriminated applicants (defined by their migration status, their residence or their gender) are specific in firms that entered the experiment. Table 2.10 shows the estimation results, using different set of control variables. Although the probability to be interviewed is the same for applicants to participating and non-participating firms (first column), there are significant differences across applicants' types: firms that accept to enter the experiment tend to more frequently call minority applicants for interviews, and to less frequently interview women. These differences are only marginally significant when using no controls; the significance further decreases when controlling for applicants' and firms' characteristics. Introducing job offer fixed effects has two consequences: first, the fixed effect absorbs differences across firms that are not related to the applicant's type; second, only firms with mixed pools of applicants play a role in identifying d_1 and d_2. This has little impact on the estimation of d_1. It does however lower the point estimate for d_2, suggesting that, if one restricts the comparison to firms with men and women in their applicant pools, there is no significant difference along gender lines between firms inside and outside the experiment. Results on hiring decisions (on the right hand side panel) yield a similar picture.

Overall, these results do suggest some differences: with standard nominative resumes, the chances of minority candidates tend to be higher in firms participating to the experiment; women's chances would instead be lower. Note that these differences may still be due to unobserved firm and applicant heterogeneity, rather than to difference in discrimination behavior. However, these results do call for a note of caution, as the population of firms entering the experiment is not representative of the overall population of firms: simple correlations suggest that they represent firms that are rather more favorable to minority applicants. Interestingly, they do not seem to be more favorable to women (if anything, they are actually less favorable). This echoes the findings of qualitative analysis of the experiment: some hiring officers said that they first perceived this experiment as concerning candidates of foreign origin or residing in deprived neighborhoods, but that participating to the experiment made them more aware of gender issues as well. It is therefore possible that firms self-selected themselves more according to their behavior concerning ethnic minorities and residents in deprived neighborhood, rather than according to their treatment of female applicants.

2.8 Impact of anonymous resumes on applicants

2.8.1 Overall impact

We start by analyzing the average impact of anonymous resumes on different groups of applicants, all firms taken together. In the next subsection, we investigate the heterogeneity of these effects according to the firms' characteristics.

Due to the experimental design, the impact of anonymous resumes on any subpopulation is immediately identified as the difference in mean outcomes between control and treatment individuals, within this subpopulation. However, the result of the policy is better defined as a relative impact: do anonymous resumes reduce the gap between applicants at risk of discrimination and other applicants? This question implies to start by defining a group of reference (presumably not victim of discrimination), and one or several groups that are potentially discriminated. Clearly, there is a trade-off between the advantage of looking at narrowly, well-defined groups, and the statistical precision allowed by the sample size. We conducted a variety of statistical tests (available on demand) to detect along which lines anonymous resumes have heterogeneous effects. We considered four dimensions along which anonymous resumes may have a differential impact: the applicant's gender, age, place of residence and migration status. Interacting these four dimensions yields 16 different groups, with 0, 1, 2, 3 or 4 potential stigmas. Anonymous resumes do not seem to impact applicants of different ages differently – perhaps simply because age can easily be derived from the work experience detailed in the resume. Finally, it turns out that the impact of anonymous resumes is well summarized by a parsimonious model:

$$Y_{ij} = \alpha_0 + \alpha_1 M_i + \alpha_2 F_i + \delta_0 T_j + \delta_1 T_j \times M_i + \delta_2 T_j \times F_i + \epsilon_{ij}, \qquad (2.2)$$

where Y_{ij} is an indicator variable equal to 1 if applicant i on job offer j is interviewed (or is hired), M is an indicator for being in the group of immigrants, children of immigrants and/or residents of deprived neighborhoods, F is an indicator for female applicants, and T is the indicator variable for the use of anonymous resumes on the job offer. Equation 2.2 is estimated by OLS, accounting for correlations between applicants on the same job offers using robust standard errors, clustered at the job offer level. We use sampling weights to account for the fact that some applicants were oversampled in the survey. Unless otherwise specified, the model is estimated only on applicants whose resumes were preselected by the PES before randomization, so as to ensure the comparability of applicants under the standard and nominative procedures. Among firms that entered the experiment, compliance to random assignment is nearly perfect[13], so that δ_0 is directly interpreted as the impact of anonymous resumes on the reference group (males who are neither immigrants, sons of immigrants nor residents in deprived neighborhoods), and δ_1 and δ_2 give the additional impact for immigrants, sons of immigrants or residents in deprived neighborhoods, on the one hand[14], and for women, on the other hand. In other words, δ_1 and δ_2 summarize how the gap between potentially discriminated applicants and other applicants is impacted by anonymous resumes.[15]

Table 2.11 gives a first pass on three questions: (i) Do anonymous resumes induce firms to interview more applicants, in order to compensate for the loss of

[13] 17 firms (13 treatment firms and 4 control firms) exited the experiment after the random assignment and therefore received standard resumes. Applicants to these firms are interviewed, and analyzed according to the initial random assignment.

[14] Disaggregated results for that composite group are discussed below.

[15] Note that α_1 and α_2 do not have a causal interpretation, as they may capture unobserved differences in applicants' productivity that are correlated with applicants' type.

information on applicants' type? (ii) Do anonymous resumes improve the relative changes of potentially discriminated applicants to be selected for a job interview? (iii) Do the effects of making applications anonymous persist after the applicant's type is revealed by the interview, so that hiring decisions are impacted?

The left panel displays the impact of anonymous resumes on the probability that a given applicants is interviewed. There is no evidence that firms reacted to anonymous resumes by selecting a larger pool of applicants for interview (column 1). But, as shown in column 2 (which corresponds to model 2.2), this overall stability hides lower callback rates for applicants of foreign origin and/or residents of deprived neighborhood, and higher callback rates for other applicants. As a result, the gap between majority candidates and candidates at risk of discrimination due to their migration status or their place of residence, which is small and not significant with standard resumes, *increases* significantly when anonymous resumes are used. The effect is large (a 10 percentage point increase, to be compared to average interview rates around 11%). This is the first key, counter-intuitive result of the experiment: overall, anonymous resumes redistribute chances to be called for a job interview, but this redistribution occurs at the expense of those that the law expected to benefit. By contrast, anonymous resumes do not significantly impact the relative chances of women to be interviewed (the point estimate is positive, but far from conventional significance levels). As expected with random assignment, these results are robust to the addition of control variables (column 3). However, adding job offer fixed effects does change the picture. The negative impact on candidates of foreign origin or residents in deprived neighborhoods is confirmed (though it is not statistically significant anymore due to a loss of precision), but anonymous resumes now seem to improve the relative chances of women: the point estimate is large, so that the effect, even if it is imprecisely estimated, is significant at the 10% level. To interpret this pattern, note that fixed effects restrict the source of variation identifying δ_2 to comparisons of male and female applicants in a given job offer – this is perhaps the most natural way to analyze the impact of anonymous resumes: how does it change the relative chances of candidates when they are competing on the same offers? In the survey of applicants we use, however, only 117 job offers (out of 598) have applicants of both genders. Part of this is due to sampling (there are applicants of both genders, but all were not interviewed); part of it, however, reflects the fact that the PES preselected only men for 31% of the experimental job offers, and only women for 17% of the others. On these job offers, anonymous resumes should not impact the relative chances of women, except if firms react to uncertainty by interviewing more candidates, and if they do so differently depending on whether they have male or female applicants. In other words, the estimate of δ_2 in columns 2 and 3 is the mean of the (presumably null) impact of anonymous resumes on job offers where the two genders do not compete, and of its impact on job offers where they indeed compete. The estimates in column 4 suggest that, in the latter job offers, woman have lower interview rates than men when resumes are nominative (a -11.1 pp difference, not statistically significant), and that anonymous resumes are effective in closing the gap. Table 2.20 checks this interpretation by estimating equation 2.11 only on job offers with candidates of both genders[16]: this allows us to estimate the effect of anonymous resumes using variations in 311 rather than 117 firms. The estimated δ_2 is still positive, statistically significant, even though slightly smaller (around 12 percentage points). [17]

The right panel of table 2.11 addresses the third question: do these effects persist one the firm knows the applicant's identity? Unfortunately, the statistical power to detect an impact of anonymous resumes on hiring decision is limited: the probability of being hired is small; only very large changes (when expressed in percents of the initial hiring probability) can be detected. Column 6 does suggest that the lower interview rates of applicants of foreign origin and/or residing in deprived neighborhood translate into lower chances to be hired (the relative impact is negative, but significant at 10%). This however is not robust to the inclusion of controls.

To summarize, table 2.11 yields four main key results:

1. Anonymous resumes do not induce firms to call more applicants for interviews.

2. Anonymous resumes decrease the relative chances of applicants of foreign origin and/or residing in deprived neighborhood to be interviewed for a job.

3. Anonymous resumes increase the relative chances of women on job offers for which they are competing with men; however, there is only half of such offers in our sample, so that this translates only in a modest and non significant change in the overall population.

4. Evidence on whether these effects on the selection of resumes by firms translate into hiring decisions is not conclusive.

Robustness checks

We performed a variety of tests to check the robustness of these key facts to potential threats to the experimental design, as well as to alternative measurements and model specifications.

A John Henry effect?

As noted below, a possible issue with the experimental design is the fact that participating firms knew that they were part of an experiment. This in itself could affect their behavior. The risk is particularly acute for control firms: they know they were observed, they also know the identity of applicants, and are therefore directly confronted to the question of calling for interviews potentially discriminated

[16] As characterized by the total pool of candidates from the administrative data, rather than the sample from the survey of applicants.

[17] We explore other dimensions of effect heterogeneity in the next subsection. The distinction between job offers with or without a mixed pools of applicants matters less when it comes to migration status or place of residence. Indeed, in our sample, 72% of job offers had applicants both from the potentially discriminated group and from the rest of the population. Restricting the estimation to these job offers does not alter the picture much (see table 2.21).

applicants. They may therefore artificially select more of these applicants in order to signal to the PES that they do not discriminate. This type of effect is known as a "John Henry" effect, by which the control group makes extra effort to perform well. Such effect could explain why treated firms appear less favorable to applicants of foreign origin and/or residing in deprived neighborhood: the negative difference would not be due to a negative impact of anonymous resumes on treatment firms, but to the positive impact of monitoring firms in the control group.

Our strategy to test for the presence of such an effect is to look at control firm hiring behavior *after* the experiment. The idea is the following: randomization ensures that control and treatment firms are comparable. But *during the experiment*, control firms change their behavior, so that they are not a valid counterfactual. However, unless being part once of an experiment (without being treated) has surprisingly lasting effects, control firm behavior should not be distorted anymore after the experiment. We therefore ran the survey of applicants on a subsample of applicants to 148 job offers that were posted by control firms after the experiment.[18] The test for the presence of a John Henry effect is very simple: we ask whether interview and hiring gaps between applicants of different groups were different before and after the experiment. Formally, we use the same type of model as above:

$$Y_{ij} = \alpha_0 + \alpha_1 M_i + \alpha_2 F_i + \delta_0 EXP_j + \delta_1 EXP_j \times M_i + \delta_2 EXP_j \times F_i + \epsilon_{ij}, \quad (2.3)$$

where Y_{ij} is an indicator variable equal to 1 if applicant i on job offer j is interviewed (or is hired), M is an indicator for being in the group of immigrants, children of immigrants and/or residents of deprived neighborhoods, F is an indicator for female applicants; last EXP is the indicator variable for job offers that were part of the experiment. The model is estimated on 807 applicants, who applied to 148 pairs of job offers posted by control firms, one job offer being included in the experiment, the other being excluded. δ_0, δ_1 and δ_2 summarize how interview and hiring rates evolved for different group of applicants in the same firms, before and after the experiment.

Table 2.12 shows no evidence of a John Henry effect. Interaction coefficients are small, and far from being significant. If anything, control firms were more rather than less favorable to applicants of foreign origin and/or residing in a deprived neighborhood when the experiment stopped. Of course, a John Henry effect may persist over time (once firms know they have been observed, they are durably more cautious); but one would expect the effect to decay. This is not the case.[19]

Insufficient anonymization?

[18] Note that relying on applicants for information on interview and hiring decisions made by the firm removes any concern that firms become aware that we continue observing them.

[19] Instead of focusing on job offers posted by control firms after the experiment, one could have looked at job offers posted before the experiment, that would be fully exempt from any John Henry effect. However, this turns out not feasible. Indeed, administrative information being insufficient, we would need to run survey of applicants on these past job offers. By construction, these surveys would often occur with a significant delay – to identify control firms, one needs to wait for them to enter into the experiment! – which would create memory bias among respondents.

As noted above, anonymization in the experiment is imperfect, as the content of resumes is not standardized. In particular, foreign background can be inferred from languages skills: can this explain why anonymous resumes did not have more favorable effects on applicants of foreign origin?

In our sample, 90% of the resumes have been processed manually. For those resumes we have information about the language skills of the candidate; in particular, we know whether the applicants speaks Arabic, or any other foreign language than those typically taught in French schools (English, German and Spanish). Let us define as "foreign" any other language than French, English, German or Spanish. When recruiters read in a candidate's resume that she has foreign language skills, they can infer that she is immigrant or child of immigrant: this will be a bad guess in only 20% of the cases. Observing no language skills in the resume is less informative: among the candidates who do not state any foreign language skills, around one third are immigrants or children of immigrants. All in all, using the languages skills as a proxy for foreign background is a successful strategy in 70% of the resumes. Focusing on Arabic, language is even a better proxy for foreign background: when recruiters read in a candidate's resume that she has Arabic language skills, they can infer that she is immigrant or child of immigrant from the Maghreb, and this will be a bad guess in only 7% of cases. Again, observing no Arabic language skills in the resume is less informative. Among the candidates who do not state Arabic language skills, 13% are immigrants or children of immigrants. All in all, Arabic language skills is a good proxy for foreign origin in 87% of resumes.

Foreign language skills are therefore a strong signal of foreign background. One may therefore suspect that the impact of anonymous resumes is lower on applicants with such skills on their resumes. We estimate the following heterogeneous treatment effect model:

$$Y_{ij} = \alpha_0 + \alpha_1 D_i \times (1 - L_i) + \alpha_2 D_i \times L_i + \delta_0 T_j + \delta_1 D_i \times (1 - L_i) \times T_j + \delta_2 D_i \times L_i \times T_j,$$

where L indicates whether the candidate states a foreign language skill, D indicates whether the candidate is potentially discriminated on basis of her foreign background and T indicates that the job offer was processed with anonymous resumes.

Results on the interview rate are displayed in table 2.13. In the first two columns, potentially discriminated candidates are defined in the usual way : they have foreign background or they live in a deprived neighborhood. In the third column, foreign background is restricted to immigrants or children of immigrants from the Maghreb. In the first column, foreign language is defined broadly (any language different from English, German or Spanish). In the last two, it is restricted to Arabic language. According to the first column, being potentially discriminated or speaking a foreign language does not affect the interview rate when resumes are nominative. Anonymization has no significant effect on non discriminated candidates (δ_0). The typical negative relative effect of anonymization on potentially discriminated candidates is estimated for both groups of candidates (δ_1 and δ_2). Effects are not heterogeneous depending on language skills: the difference between the two coefficients (around 3 percentage points) is not statistically significant. The second and third columns confirm the absence of heterogeneous effects. The analysis of hiring rates leads to the same conclusion. We discuss possible interpretations of this finding below.

Alternative measures of applicants' background

The applicant's background enters model 2.2 in a quite specific way, imposing the same effect for being of foreign background (immigrant or child of immigrant, denoted by the indicator variable I), residing in deprived neighborhood (denoted by the indicator variable Z), and cumulating the two characteristics. Other models are possible: for instance, the impact of potentially discriminated characteristics may cumulate (implying an additive model, with I and Z entering separately); they may reinforce each other (implying a model with I, Z and $Z \times I$), etc. The correct specification is an empirical question. Moreover, it is not obvious how foreign background should be measured. Names are directly impacted by anonymous resumes, so that this may be the relevant measures. However, coding whether family names denote a foreign background is not immediate. Moreover, even if first names and surnames do not denote a foreign background, a picture ID might. In that case, measuring foreign background with the applicant's migration status may be more relevant. Again, the appropriate measure is an empirical question.

Table 2.22 displays alternative possible specifications. Looking at coefficients on $T \times I$, $T \times Z$ and $T \times I \times Z$ in columns 2, 4, 6 and 8 shows that the effects of I and Z do not cumulate: basically, having only one of the two characteristics or the two of them does not modify the (negative) impact of T. This is why our preferred specification characterizes applicants as potentially discriminated due to their background when they are either of foreign background, or residing in deprived neighborhood, or both. In columns 3 to 8, three possible alternative measures of foreign background – being an immigrant, or being the child of an immigrant, or having a Muslim or African-sounding first name – yield similar results to our preferred measure (which groups immigrants and children of immigrants). Point estimates, however, tend to be lower, and the effect is no longer statistically significant when considering only Muslim or African-sounding name. A plausible explanation for that is attenuation bias due to measurement error. For instance, we know from table 2.4 that about 40% of applicants with a foreign background are not signalled by a Muslim or African-sounding name. If they are actually detected by firms, this contaminates the group of reference, creating a downward bias on the coefficient of interest. Overall, table 2.22 justifies model 2.2 as a parsimonious but appropriate way to model the differential impact of anonymous resumes.

Other specification issues

Table 2.23 displays additional robustness checks. First, we check whether sampling weights make a difference. The coefficient on $T \times M$ becomes smaller and marginally significant only. This may be due the fact that, among applicants from a foreign background, applicants with a Muslim or African-sounding name have been oversampled (this was the only information on foreign background available at the time of sampling). The lower point estimate suggests that the negative effect of anonymous resumes could be smaller on that group. The difference, however, is far from significant. In column 3, we check whether expanding the sample to applicants whose resumes were pre-selected by the PES agents after randomization makes a difference. Again, the coefficient on $T \times M$ is smaller. There remains however a suspicion that the pools of candidates in the treatment and control group are no longer comparable. Last, we check that using a logit model rather than a linear probability specification does not affect the results.

2.8.2 Heterogeneous effects

An important question is whether the main effects summarized in table 2.11 apply generally, or whether anonymous resumes impact the gap between different groups differently on different subpopulations of jobs, applicants or firms.

The obvious problem here is the curse of dimensionality (what we want to analyze here is a difference in differences in differences: anonymous vs standard resumes, for applicants with or without potentially discriminated characteristics, in subpopulations A and B), and the corresponding risk is data mining. We did however replicate the analysis of table 2.11 on different subsamples defined by the job skill level, the industry, the applicant's education level, whether the firm reported HR policies against discriminations or not, whether the firm finds it difficult to hire[20], etc. No systematic and significant differences appeared. A better approach is certainly to start from priors on dimensions of heterogeneity that should matter, from a theoretical perspective. We consider two of them.

Labor market segmentation by gender

The first dimension to consider relates to a labor market segmentation hypothesis: if there are men jobs, women jobs, and jobs for men and women, one should not expect anonymous resumes to impact these jobs similarly. The most likely predictions is that anonymous resumes will not change the prospects of women for women jobs (that they will get anyway), nor for men jobs (which they will not get), but that they may improve their chances on jobs for which men and women are competing. As discussed above, the contrast between columns 3 and 4 of table 2.11 tends to confirm this hypothesis. More precisely, tables 2.11 and 2.20 show that anonymous resumes have no impact on job offers for which PES agents select only men or only women, but that anonymous resumes positively impact women when the PES agents preselect a mixed pool of applications. The question that remains open is whether the PES agents' pre-selection reflect a feature of the labor market (segmentation). To check this, we analyze the share of female job-seekers by type of job thought.[21] The distribution of the share of female job-seekers across jobs is displayed on figure 2.2. We define 3 types of job sought:

[20] In the firm survey, there is very detailed information about the recruitment process in the experiment (type of interviews, number of persons in charge of interviews...), and much less about usual recruitment practices. Because treatment may have affected the recruitment process in the experiment, we do not consider heterogenous effect along this detailed information. The recruitment process is considered as an outcome in the last section.

[21] This analysis uses an additional data source, the administrative files kept by the PES on all registered job-seekers.

1. Male dominated jobs: when the share of female unemployed seeking this type of job is less than 25% (one example of position is security guard))

2. Women dominated jobs : when the share of female unemployed seeking this type of job is more than 75%(for instance, secretary)

3. Mixed jobs (the complement)

Among the stock of registered unemployed, those jobs represent respectively 36%, 14% and 50%. This indicator of gender segregation is a good predictor of the segregation observed in our sample displayed on figure 2.1: the coefficient of correlation is 0.72. 60% of jobs offers have both indicator consistent. For example, 65% of the job offers predicted as mixed by the external segregation measure are indeed mixed in our sample.

Overall, the analysis confirms that anonymous resumes improve the chances of women to be interviewed on jobs for which labor supply is mixed. However, the French labor market features persistent segmentation, so that some positions only attract women's applications, whereas other only attract men's applications. As expected, we see no impact of anonymous resumes on these segments. One possibility though is that anonymous resumes could, in the long run, have a "calling" effect: if women feel they now have a fair chance to get positions that used to be "men's jobs" thanks to anonymous resumes, they may start competing for these positions too. Such effect is absent from our evaluation, where applicants were most likely not aware of the use of anonymous resumes.

Homophily

The second hypothesis is known in the literature as the homophily hypothesis: in our setting, individuals would tend to discriminate against members who do not belong to their own group. With this hypothesis in mind, we made specific effort to characterize the group of the recruiters in experimental firms (see the data section). Table 2.14 (respectively, 2.15) estimates equation 2.2 after stratifying the sample of recruiter by gender (respectively, according to her network).

Table 2.14 shows a pattern that is consistent with the homophily hypothesis. Male recruiters tend to select fewer women for interview, and to hire fewer of them, while female recruiters tend to select fewer men (as shown by columns 3 and 6, the differences are significant at the 5% level). Of course, alternative interpretations are possible, as the recruiter gender may be correlated with other characteristics of the firm.[22]. Turning to interaction effects, we find that anonymous resumes undo this differential treatment: the interaction coefficient on T $\times$ woman is positive when the hiring recruiter is a man, negative when it is a woman. This difference is strongly significant. In other words, anonymous resumes counteract the tendency of hiring officers to select applicants of their own gender: it therefore equalizes the chances of men and women, independently from the gender of the recruiter. Most interestingly,

[22]We do not however find that it is correlated with the fact that the applicants' pool is mixed or not

this has consequences on the final recruitment decision, after the hiring officer has actually met the candidate.

Table 2.15 looks for a similar pattern for applicants of foreign background: are they treated differently depending on the background of the recruiter? There are unfortunately very few recruiters with a foreign background in our sample. A more useful measure is provided by asking the recruiter about the first names of her friends: this allows to identify recruiters who cite at least one African or Muslim-sounding name among three friends. We do not find evidence of differential treatment with standard resumes; correspondingly, we do not find that anonymous resumes affect applicants with a foreign background differently depending on the identity of the hiring officer.

2.9 Impact of anonymous resumes from the recruiter perspective

We now evaluate the effects of anonymous resumes on the costs of the recruitment process.[23] During the experiment, the direct costs of anonymization have been paid by the Public Employment Agency. We thus focus on more indirect but no less important costs, such as the number of interviews, the time to recruit... We expect those costs to increase with anonymization. By reducing the level of information in the first stage of the recruitment process, firms may report their selections to further stages, and increase the number of interviews or tests, which are typically more costly. Those modifications of the recruitment process may also delay the hiring date, increasing the opportunity cost of keeping a job unfilled.

A particular concern is whether anonymization affects match quality, as measured by wages or output. A direct measure of output is not available, but we take as a proxy whether the trial period was successful, which should reflect that output is above a minimum threshold. We also estimate the effect of anonymization on the hiring wage. Note however that hiring wages do not only reflect productivity, but also the outside labor market options of the candidate. Assume that anonymization does not affect the productivity of the hired candidate but that hired candidates are more often from the discriminated minority group. Wages may still decrease as a result of the candidate's lower bargaining power.

2.9.1 Crowding out effects

Before performing the cost benefit analysis from the recruiter perspective, we estimate possible crowding out effects of candidates from the Public Employment Service. As a response to the lower level of information on candidates sent by the PES, firms may activate other more costly channels to meet candidates.

[23] We also considered whether the costs and the nature of the recruitment process differs between firms that entered the experiment and firms that did not enter. Except if noted otherwise, we found no significant difference (results omitted).

Around one out of two applications received by the recruiter come from the PES and one out of four interviewed candidates are sent by the PES (line 1 and 2,column 1 in table 2.16). Last, one out of three hired candidates are sent by the PES. This highlights the fact that recruiters do not rely exclusively on the PES to drain candidates. However small, the share of PES candidates does not decrease with anonymous resumes: there is no evidence that anonymous resumes in the PES leads to a crowding out of the candidates it sends.

2.9.2 Costs

Anonymous resumes have not altered the probability of successful recruitment. Around four out of five hirings were completed at the time of the survey (see line 3 column 1 of table 2.17) and the difference between control and test (column 2) is small and not significant. Anonymous resumes have not altered the probability that the recruitment had been stopped without any hiring (line 1). The mean time to hiring is 49 days in the control group. The first and third quartiles of the distribution are 20 and 72 days. Anonymous resumes do not alter that distribution.

Overall, these findings suggest that anonymous resumes do not increase the costs associated with foregone output due to unsuccessful or delayed hiring.

We now turn to the hiring process itself. Half of the recruiters in the control group receive at most 12 applications and interview at most 6 candidates (line 1 and 2, column 1 in table 2.18). The median numbers of applications and interviews is not affected by the use of anonymous resumes.

Recruiters select candidates thanks to various tools : phone interviews, collective, individual interviews and tests (in situ). Individual interviews are conducted by four out of five recruiters, phone interviews by two out of five recruiters, tests by one out of five recruiters (line 3, 4 and 5; column 1 in table 2.18). Collective interviews are relatively marginal. Anonymous resumes do not lead recruiters to change their mix of selection tools (column 2). The mean number of tools used is 1.6 in both control and test group (line 6).[24]

We find no increase in the number of recruiters or the total working time devoted to the recruitment. In the control group, around 2 recruiters take part to the process and half of the job offers are filled in less than 8 hours and a half.

2.9.3 Benefits

Even in the absence of cost increase, it is relevant from the recruiters' point of view to estimate potential benefits associated with anonymous resumes. In table 2.19, we analyze hired candidates as described by the recruiters in the firm survey. Note that most of the hired candidates were not addressed by the PES: in this subsection we

[24]Recruiters who withdraw before randomization tend to have a significantly larger selection toolkit (1.8 mean number of tools). Again this shows that their recruitment process is more intensive, leading presumably to a faster recruitment (results not presented here).

analyze a broader population than in the previous section on candidates. Indeed, from the recruiter perspective, this global effect is the relevant one.

Four hired candidates out of five successfully complete their trial period. Recruiters are generally satisfied with the first tasks performed by the hired candidate or more generally with his/her adequation to the job. Moreover, match quality as measured by successful trial period or recruiters' subjective satisfaction is not affected by the use of anonymous resumes. One hired candidate out of five is paid the minimum wage. Half of the workers who are paid more than the minimum wage earn more than 1 715 euros per month (gross wage without any bonuses). The wage distribution is concentrated just above the minimum wage (1 350 euros). The first and third quartiles are respectively 1.1 and 1.63 of the minimum wage. Anonymous resumes do not affect the share of hired candidates paid the minimum wage, nor the median or first quartile of the wage distribution. The third quartile is significantly lower by 250 euros. This latter result, however, is not robust to the addition of controls; moreover, it is not clear whether such an effect should be interpreted in terms of productivity or bargaining power.

Overall, we find no evidence that anonymous resumes change hiring costs, labor costs and match quality. Two caveats must be kept in mind, though. First, the PES took in charge the anonymization procedure itself and these costs are not included here; second, we only test for short-term effects for filling one position: anonymous resumes may, in the long run, lead firms to modify more substantially their hiring process.

Table 2.1: Sample size and response rates in the applicant survey

	Population size (a)	# sampled for survey (b)	# of respondents (c)	Sampling rate (b/a)	Response rate (c/b)	Information rate (c/a)
Control	3443	1520	1012	0.44	0.67	0.29
Treatment	3299	1464	945	0.44	0.65	0.29
At risk of discrimination	2312	1369	900	0.59	0.66	0.39
Other applicants	4430	1615	1057	0.36	0.65	0.24

Note : This table displays population and sample size among applicants to job offers entering the experiment. The first two lines distinguish applicants according to whether resumes were anonymous (treatment) or not (control); the last two lines distinguish applicants who have an African / Muslim-sounding name or live in deprived neighborhoods from other applicants.

Table 2.2: Sample size and response rates in the firm survey

	Population size (a)	# sampled for survey (b)	# of respondents (c)	Sampling rate (b/a)	Response rate (c/b)	Information rate (c/a)
Control	385	385	229	100.0	59.5	59.5
Treatment	366	366	212	100.0	57.9	57.9
Withdrew before randomization	254	254	134	100.0	52.8	52.8
Refused the experiment	608	335	146	55.1	43.6	24.0
Not invited	4714	542	281	11.5	51.8	6.0

Note: This table displays population and sample size among job offers eligible for the experiment (one job offer per plant). The first two lines display job offers handled with standard resumes (control job offers) or anonymous resumes (treatment job offers). The third line corresponds to firms who accepted to enter the experiment but withdrew of filled their position before the PES had provided a first lot of resumes and randomization could take place. The last two lines correspond to plants that refused the experiments or that were not invited to participate, despite the fact they were eligible.

Table 2.3: Measures of risk of discrimination

Discriminatory characteristic	Mean	Correlation with							
		(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
Woman (a)	49	1							
Age below 30 (b)	48	0.15	1						
Age above 50 (c)	11	-0.07		1					
Deprived neighborhood (d)	24	-0.09	0.09	0.09	1				
African or Muslim-sounding name (e)	23	-0.12	0.05	-0.02	0.15	1			
Immigrant (f)	22	-0.10	-0.03	0.00	0.14	0.38	1		
Son of immigrant (g)	16	-0.02	0.14	-0.10	0.11	0.36	-0.24	1	
Immigrant or son of immigrant (h)	38	-0.10	0.08	-0.07	0.20	0.59	0.67	0.56	1

Source : Applicants' main sample; 1260 observations.

Table 2.4: Immigrants and applicants with African or Muslim-sounding name

	Immigrant or son of immigrant	Other applicant	Total
African or Muslim-sounding name	21%	2%	23%
Other name	17%	60%	77%
Total	38%	62%	100%

Note : This table compares the risk of discrimination due to migratory origin based on two indicators: having an African or Muslim-sounding name, or being an immigrant or the child of an immigrant. Each cell gives the frequency within the experimental population (using sampling weights).

Table 2.5: Applicants' characteristics according to gender, neighborhood of residence and migration status

	Intercept	Women	Deprived neighborhood	Immigrant	Child of immigrant
Years of education	12.684***	.418**	-.518***	.653***	-.170
	(.144)	(.169)	(.195)	(.245)	(.209)
Age	35.263***	-2.542***	-1.387***	.283	-3.505***
	(.410)	(.454)	(.505)	(.553)	(.559)
Work experience (in years)	4.633***	-.614***	-.417**	-.381*	-.467**
	(.150)	(.169)	(.183)	(.204)	(.219)
Experience on a similar job (in years)	2.265***	-.555***	-.013	-.807***	-.789***
	(.185)	(.187)	(.229)	(.220)	(.234)
Has been looking for a job for more than a year	.136***	-.035**	.027	-.013	-.026
	(.013)	(.015)	(.017)	(.019)	(.019)
(.025)	(.028)	(.030)			
Reservation wage is minimum wage	.422***	.155***	.102***	.041	.063**
	(.019)	(.022)	(.025)	(.028)	(.029)
Reservation wage (euros)	1841.086***	-203.178***	-134.363***	-30.362	-93.493***
	(26.912)	(25.787)	(25.786)	(31.380)	(35.377)
Has a driving license	.700***	-.130***	-.011	-.153***	-.095***
	(.020)	(.023)	(.025)	(.029)	(.031)
Speaks a non-European language	.039***	.008	-.019	.306***	.214***
	(.011)	(.016)	(.018)	(.024)	(.024)

Note : Each line corresponds to an OLS regression of an applicant credential on indicator variables for women, residents in deprived neighborhood, immigrants and children of immigrants. Standard errors in brackets. Significant differences at 10%, 5% and 1% are denoted by *, **, and ***. Sample: all applicants from applicant survey (1,957 observations). Source: applicant survey, except for driving license and language skills, directly coded from resumes.

Table 2.6: Balancing between control and treated candidates

	Test	Control
Candidates' characteristics		
Women	0.51	0.47
Less than 26	0.29	0.29
More than 50	**0.13**	0.09
Deprived neighborhood (1)	0.23	0.25
Immigrant (2)	0.22	**0.23**
Child of immigrant (3)	0.18	0.15
(1), (2) or (3)	0.50	0.49
African or Muslim-sounding name (4)	0.24	0.24
More than 12 years of education	0.72	0.75
Number of candidates sent to recruiter		
candidates (1) or (4)	2.08	2.05
other candidates	3.45	3.42

Source : Candidates' survey

Table 2.7: Comparison of firms and positions offered according to participation in the experiment

	Control (a)	Treatment (b)	Withdrew before randomization (c)	Refused to participate (d)	Not invited to participate (e)	Test of difference (b) vs (a)	(c) vs (a)	(d) vs (a)	(e) vs (a)
Firm with less than 100 employees	34.4	32.4	34.6	35.8	30.4				
Firm with 100 to 200 employees	16.8	17.4	18.4	18.8	17.8				
Firm with more than 200 employees	48.8	50.2	47	45.4	51.8				
Non-merchant services	24.7	23.2	28	21.5	30.2				**
Merchant services	47	47	49.2	49.8	47.3				
Manufacturing	13.8	16.9	9.8	11.5	8.5				***
Construction	3.4	3.6	2.4	2.8	3.4				
Upper occupations	9.9	6.3	5.5	4.9	5.4	*	**	***	***
Intermediary occupations	24.4	26	27.6	20.9	21.1				
Skilled white or blue collar	55.3	58.7	52.4	63.7	58.8			***	
Unskilled white or blue collar	10.4	9	14.6	10.5	14.7				***
Indefinite duration contract	66.5	62.6	63.4	62.7	59.5				***
Contract for more than 6 months	86	82.2	79.1	82.9	83.3				
Number of observations	385	366	254	608	4719				

Note : Significant differences at 10%, 5% and 1% are denoted by *, **, and ***. **Source:** PES administrative data (SAGE, TCD-banque de CV, MER). All eligible job offers.

Table 2.8: Comparison of firms and positions offered according to participation in the experiment (2)

	Control (a)	Treatment (b)	Withdrew before randomization (c)	Refused to participate (d)	Not invited to participate (e)	Test of difference (b) vs (a)	(c) vs (a)	(d) vs (a)	(e) vs (a)
Firm's characteristics									
International group	33.2	40.3	25.0	34.2	31.4				
Firm has led actions against discriminations	58.0	48.7	50.0	47.5	47.7	*		*	**
Firm has an employee in charge of fighting discriminations	28.2	24.1	21.9	24.5	28.7				
Characteristics of a typical recruitment									
Easy	68.4	63.3	59.2	57.5	62.5		*	**	
Through the PES	79.4	80.5	84.6	89.0	77.1			**	
Uses PES selection of resume	47.0	45.0	41.8	46.1	55.3				*

Note : Significant differences at 10%, 5% and 1% are denoted by *, **, and ***. **Source:** Firm survey.

Table 2.9: Comparison of firms and positions offered according to participation in the experiment (3)

	Control (a)	Treatment (b)	Withdrew before randomization (c)	Refused to participate (d)	Not invited to participate (e)	Test of difference (b) vs (a)	(c) vs (a)	(d) vs (a)	(e) vs (a)
Job characteristics									
Involves teamwork	85.2	75.4	81.2	83.0	84.2	**			
Frequent customer contact	71.9	67.9	75.2	70.4	75.7				
Hiring officer characteristics									
Woman	63.8	57.8	63.6	65.8	56.4				*
College graduate	59.0	62.1	62.9	53.4	58.6				
Age	40.6	39.5	39.8	40.5	41.2				
Firm tenure (in years)	9.1	8.0	6.9	8.0	8.3		**		
Experience in hiring (in years)	8.5	9.2	9.6	9.2	10.4				***
French as mother tongue	97.8	97.6	91.7	96.5	95.2		**		
Immigrant	2.6	2.4	4.5	3.4	2.9				
Immigrant or daughter of immigrant	11.4	10.0	10.6	11.6	11.8				
At least one friend (out of 5) with Muslim or Afr. name	24.6	22.0	30.2	26.5	24.7				
At least one colleague (out of 5) with Muslim or Afr. name	27.5	27.0	27.0	26.6	29.4				

Note : Significant differences at 10%, 5% and 1% are denoted by *, **, and ***. **Source:** Firm survey.

Table 2.10: Comparison of interview and hiring rates in firms entering the experiment (with anonymous resumes) and non-participating firms

	interview				Recruitment			
Intercept	0.132***	0.103***	0.097***		0.059***	0.038**	0.034**	
	(0.025)	(0.031)	(0.032)		(0.015)	(0.015)	(0.016)	
Deprived neighborhood		-0.077**	-0.070*	-0.064*		-0.035	-0.034	-0.055**
or foreign origin		(0.038)	(0.037)	(0.036)		(0.027)	(0.027)	(0.027)
Woman		0.122***	0.116***	0.022		0.070***	0.068***	0.011
		(0.041)	(0.040)	(0.036)		(0.026)	(0.025)	(0.028)
Entered the experiment (P)	0.011	0.030	0.044		-0.017	0.002	0.009	
	(0.030)	(0.042)	(0.046)		(0.018)	(0.023)	(0.025)	
P × deprived neighborhood		0.083*	0.071	0.072		0.027	0.025	0.052
or foreign origin		(0.048)	(0.050)	(0.050)		(0.032)	(0.033)	(0.035)
P × woman		-0.105**	-0.096*	-0.042		-0.058*	-0.052*	-0.012
		(0.052)	(0.051)	(0.062)		(0.031)	(0.031)	(0.050)
Controls	No	No	Yes	Yes	No	No	Yes	Yes
Job offer effects	No	No	No	Yes	No	No	No	Yes
Observations	1,688	1,688	1,688	1,688	1,688	1,688	1,688	1,688
Job offers	631	631	631	631	631	631	631	631
Plants	365	365	365	365	365	365	365	365

Note : each column corresponds to one regression. Robust standard errors in parentheses. Significant differences at 10%, 5% and 1% are denoted by *, **, and ***. **Source:** survey of applicants.

Table 2.11: Impact of anonymous resumes

	Interview				Recruitment			
Intercept	0.106***	0.131***	0.129***		0.022***	0.016**	0.022**	
	(0.016)	(0.035)	(0.034)		(0.006)	(0.007)	(0.009)	
Deprived neighborhood or foreign origin		-0.028	-0.010	0.023		0.002	0.001	0.016
		(0.032)	(0.032)	(0.050)		(0.012)	(0.013)	(0.025)
Woman		-0.022	-0.042	-0.111		0.011	-0.002	0.003
		(0.033)	(0.035)	(0.086)		(0.012)	(0.014)	(0.044)
Anonymous resume (T)	0.005	0.042	0.037		0.013	0.025*	0.023*	
	(0.023)	(0.047)	(0.044)		(0.011)	(0.015)	(0.014)	
T × deprived neighborhood or foreign origin		-0.100**	-0.090**	-0.117		-0.035*	-0.025	-0.046
		(0.044)	(0.044)	(0.074)		(0.020)	(0.020)	(0.035)
T × woman		0.028	0.039	0.201*		0.009	0.006	-0.001
		(0.045)	(0.043)	(0.109)		(0.021)	(0.020)	(0.045)
Controls	No	No	Yes	Yes	No	No	Yes	Yes
Job offer effects	No	No	No	Yes	No	No	No	Yes
Observations	1,260	1,260	1,260	1,260	1,260	1,260	1,260	1,260
Job offers	598	598	598	598	598	598	598	598
R-squared	0.109	0.128	0.173	0.657	0.030	0.037	0.082	0.582

Note : Robust standard errors in parentheses. Significant differences at 10%, 5% and 1% are denoted by *, **, and ***. **Source:** survey of applicants.

Table 2.12: Interview and recruitment rates in control firms before and after the experiment: test of a John Henry effect

	Interview			Recruitment		
Intercept	0.156***	0.141***	0.102*	0.053***	0.042	0.010
	(0.025)	(0.045)	(0.060)	(0.015)	(0.030)	(0.034)
Deprived neighborhood		0.020	0.017		-0.002	-0.006
or foreign origine		(0.051)	(0.052)		(0.030)	(0.031)
Woman		0.010	0.012		0.026	0.032
		(0.051)	(0.052)		(0.031)	(0.033)
Experimental job offer (EXP)	-0.025	-0.019	0.083	-0.020	0.004	0.078
	(0.033)	(0.055)	(0.113)	(0.019)	(0.035)	(0.054)
EXP × deprived neighborhood		-0.031	-0.038		-0.023	-0.009
or foreign origin		(0.064)	(0.067)		(0.034)	(0.035)
EXP × woman		0.020	0.021		-0.032	-0.034
		(0.070)	(0.070)		(0.038)	(0.036)
Controls	No	No	Yes	No	No	Yes
Observations	807	807	807	807	807	807
Job offers	296	296	296	296	296	296
R-squared	0.147	0.148	0.168	0.047	0.051	0.077

Note : Robust standard errors in parentheses. Significant differences at 10%, 5% and 1% are denoted by *, **, and ***. **Source:** survey of applicants.

Table 2.13: Heterogenous effect on the interview rates : Insufficient anonymization ?

Origin	Foreign	Foreign	Maghreb
Language	Foreign	Arabic	Arabic
	(1)	(2)	(3)
Intercept	.127***	.127***	.124***
	(.029)	(.029)	(.025)
Deprived neighborhood or foreign origin	-.033	-.030	-.032
No foreign language skills in the resumes	(.036)	(.035)	(.033)
Deprived neighborhood or foreign origin	-.039	-.058	-.065
Foreign language skills in the resumes	(.044)	(.049)	(.046)
Anonymous Resumes (T)	.051	.051	.024
	(.044)	(.044)	(.036)
T x deprived neighborhood or foreign origin	-.090*	-.097**	-.064
No foreign language skills in the resumes	(.050)	(.049)	(.044)
T x deprived neighborhood or foreign origin	-.126**	-.106*	-.069
Foreign language skills in the resumes	(.055)	(.060)	(.054)

Source: survey of applicants. Linear probability model with sampling weights. No controls. No fixed effects. **Note** : Robust standard errors in parentheses. Significant differences at 10%, 5% and 1% are denoted by *, **, and ***.

Table 2.14: Homophily in gender

	Interview			Recruitment		
	Male recruiter	Female recruiter	test	Male recruiter	Female recruiter	test
Intercept	0.225**	0.064**		0.025	0.023*	
	(0.097)	(0.027)		(0.019)	(0.013)	
Deprived neighborhood	-0.062	-0.019		0.010	-0.028	
or foreign origin	(0.089)	(0.034)		(0.025)	(0.022)	
Woman	-0.155**	0.050	**	-0.024	0.040*	**
	(0.073)	(0.037)		(0.020)	(0.021)	
Anonymous resumes (T)	-0.032	0.207***	**	0.025	0.018	
	(0.112)	(0.073)		(0.031)	(0.030)	
T x deprived neighborhood	-0.126	-0.123*		-0.065*	0.030	*
or foreign origin	(0.105)	(0.070)		(0.037)	(0.033)	
T x woman	0.220**	-0.175**	***	0.063*	-0.076**	***
	(0.092)	(0.075)		(0.034)	(0.033)	
Observations	289	436		289	436	
R-squared	0.193	0.145		0.054	0.047	

Source: survey of applicants. Linear probability model. No controls. No fixed effects. **Note** : Robust standard errors in parentheses. Significant differences at 10%, 5% and 1% are denoted by *, **, and ***.

Table 2.15: Homophily in origin

	Interview			Recruitment		
	Has the recruiter foreign friends ?			Has the recruiter foreign friends ?		
	No	Yes	test	No	Yes	test
Intercept	0.134**	0.093*		0.037**	0.027	
	(0.058)	(0.054)		(0.016)	(0.028)	
Deprived neighborhood	-0.035	0.009		-0.027	-0.026	
or foreign origin	(0.054)	(0.059)		(0.027)	(0.032)	
Woman	-0.005	-0.046		0.022	0.024	
	(0.058)	(0.064)		(0.029)	(0.028)	
Anonymous resumes (T)	0.069	0.167		-0.005	0.005	
	(0.081)	(0.122)		(0.024)	(0.036)	
T x deprived neighborhood	-0.123	-0.236**		0.025	-0.023	
or foreign origin	(0.079)	(0.108)		(0.033)	(0.047)	
T x woman	0.007	0.001		-0.039	0.022	
	(0.080)	(0.108)		(0.034)	(0.043)	
Observations	425	159		425	159	
R-squared	0.148	0.168		0.037	0.061	

Source: survey of applicants. Linear probability model. No controls. No fixed effects. **Note** : Robust standard errors in parentheses. Significant differences at 10%, 5% and 1% are denoted by *, **, and ***.

Table 2.16: Crowding out effects

	Intercept (nominative)	Anonymous resume
Share of candidates sent by the PES	.563***	.0001
	(.025)	(.038)
Share of interviewed candidates sent by the PES	.230***	.056
	(.023)	(.037)
Among successful recruitment		
Hiring from the PES	.348***	.010
	(.036)	(.052)
Nb. of successful recruitment observed	178	162

Surveyed recruiters with successful or stopped recruitment. Results are from median regressions on line 1 and 2, and from mean regressions on line 3 and 4. For the median regression, standard errors are bootstrapped.

Table 2.17: Recruitment Issue

	Intercept (nominative)	Anonymous resume
Recruitment stopped	.140***	-.022
	(.023)	(.032)
Recruitment in progress	.083***	.030
	(.018)	(.028)
Successful recruitment	.777***	-.013
	(.028)	(.040)
	Time to hiring...	
Mean in days	48.510***	-.968
	(2.792)	(4.086)
First quartile (in days)	20.000***	0.000
	(1.980)	(3.317)
Third quartile (in days)	72.000***	-9.000
	(6.253)	(8.292)
Nb. of successful recruitment observed	178	162

Recruiters responding to the survey. Linear probability model.

Table 2.18: The recruitment process in details

	Intercept (nominative)	Anonymous resume
Number of candidates (median)	12.000***	0.000
	(2.107)	(2.857)
Number of interviews (median)	4.000***	0.000
	(.597)	(.785)
Share of recruitment with ...		
Phone interviews	.399***	.044
	(.034)	(.051)
Collective interviews	.067***	.014
	(.017)	(.026)
Individual interviews	.857***	-.028
	(.024)	(.037)
Tests	.224***	.011
	(.029)	(.042)
Number of selection methods	1.581***	.031
	(.059)	(.092)
Number of recruiters involved	1.897***	.059
	(.051)	(.083)
Total time recruiters have spent for...		
Phone interviews (median in minutes)	90.000***	0.000
	(11.703)	(24.329)
Individual interviews (median in minutes)	300.000***	-60.000
	(29.821)	(42.032)
Tests (median in minutes)	120.000***	0.000
	(33.596)	(69.353)
Total (median in hours)	8.500***	.500
	(.820)	(1.108)

Source : Surveyed recruiters with successful or stopped recruitment.

Table 2.19: Match quality

(4)	Intercept (nominative) (5)	Anonymous resume
Successful trial period	.818***	-.016
	(.033)	(.048)
Recruiter's satisfaction (1-10 scale) about...		
early tasks	7.320***	.058
	(.159)	(.224)
Wage (monthly gross wage, without bonuses)		
Hired candidate paid the minimum wage	.220***	.008
	(.035)	(.051)
Median (except min wage earners)	1715.000***	-15.000
	(57.268)	(77.766)
First quartile (except min wage earners)	1500.000***	-50.000
	(24.865)	(49.592)
Third quartile (except min wage earners)	2200.000***	-250.000**
	(106.982)	(122.712)
Nb. of observations	141	127

Source : successful recruitment in the recruiters' survey for which the recruiter accept to communicate information on the hired candidate.

2.10 Appendix : complementary tables

Table 2.20: Impact of anonymous resumes: job offers with male and female applicants

	Interview			Recruitment		
Intercept	0.120***	0.185***	0.180***	0.026***	0.016	0.014
	(0.021)	(0.054)	(0.047)	(0.009)	(0.011)	(0.012)
Deprived neighborhood		-0.035	-0.013		0.007	0.007
or foreign origin		(0.043)	(0.042)		(0.017)	(0.017)
Woman		-0.086*	-0.094**		0.012	0.011
		(0.047)	(0.046)		(0.017)	(0.017)
Anonymous resumes (T)	-0.042	-0.076	-0.069	-0.013	0.001	0.012
	(0.029)	(0.065)	(0.060)	(0.010)	(0.013)	(0.015)
T × deprived neighborhood		-0.060	-0.062		-0.011	-0.008
or foreign origin		(0.057)	(0.057)		(0.019)	(0.020)
T × woman		0.119**	0.125**		-0.015	-0.025
		(0.057)	(0.055)		(0.019)	(0.022)
Observations	714	714	714	714	714	714
Job offers	311	311	311	311	311	311
R-squared	0.105	0.129	0.208	0.022	0.024	0.060

Note : Robust standard errors in parentheses. Significant differences at 10%, 5% and 1% are denoted by *, **, and ***. **Source:** survey of applicants.

Table 2.21: Impact of anonymous resumes: job offers with applicants of foreign origin or deprived neighborhood and other applicants

	Interview			Recruitment		
Intercept	0.087***	0.092***	0.097***	0.022***	0.015*	0.026**
	(0.015)	(0.034)	(0.031)	(0.007)	(0.009)	(0.011)
Deprived neighborhood		0.010	0.016		0.003	-0.001
or foreign origin		(0.031)	(0.031)		(0.013)	(0.015)
Woman		-0.023	-0.045		0.011	-0.007
		(0.031)	(0.033)		(0.013)	(0.016)
Anonymous resumes (T)	-0.001	0.050	0.044	0.006	0.024	0.020
	(0.021)	(0.048)	(0.043)	(0.012)	(0.018)	(0.017)
T × deprived neighborhood		-0.101**	-0.089**		-0.026	-0.018
or foreign origin		(0.044)	(0.044)		(0.022)	(0.023)
T × woman		0.010	0.028		-0.009	-0.004
		(0.042)	(0.041)		(0.022)	(0.022)
Observations	1,005	1,005	1,005	1,005	1,005	1,005
Job offers	439	439	439	439	439	439
R-squared	0.086	0.098	0.138	0.025	0.028	0.068

Note : Robust standard errors in parentheses. Significant differences at 10%, 5% and 1% are denoted by *, **, and ***. **Source:** survey of applicants.

Table 2.22: Robustness analysis: impact of anonymous resume by origin and neighborhood of residence // with different measures of foreign origin

	Foreign origin measured as							
	Immigrant or child of immigrant		Child of an immigrant		Immigrant		Muslim or African-sounding name	
C	0.119***	0.131***	0.111***	0.121***	0.113***	0.124***	0.119***	0.132***
	(0.026)	(0.035)	(0.022)	(0.030)	(0.023)	(0.032)	(0.023)	(0.034)
Deprived neighborhood	-0.026		-0.015		-0.018		-0.031	
or foreign background	(0.031)		(0.030)		(0.030)		(0.030)	
Deprived neighborhood		0.014		-0.011		0.010		-0.009
		(0.051)		(0.036)		(0.039)		(0.040)
Foreign background		-0.040		-0.029		-0.035		-0.062*
		(0.036)		(0.042)		(0.039)		(0.037)
Deprived neighborhood		-0.013		0.035		-0.025		0.037
and foreign background		(0.063)		(0.071)		(0.059)		(0.062)
Woman		-0.023		-0.020		-0.022		-0.026
		(0.033)		(0.032)		(0.033)		(0.034)
Anonymous resume (T)	0.057	0.039	0.034	0.014	0.033	0.013	0.027	0.004
	(0.040)	(0.047)	(0.032)	(0.039)	(0.034)	(0.042)	(0.034)	(0.043)
T × deprived neighborhood	-0.103**		-0.084**		-0.072*		-0.060	
or foreign background	(0.044)		(0.039)		(0.041)		(0.041)	
T × deprived neighborhood		-0.144**		-0.061		-0.111**		-0.068
		(0.061)		(0.047)		(0.048)		(0.051)
T × foreign background		-0.102**		-0.101**		-0.063		-0.040
		(0.048)		(0.049)		(0.051)		(0.050)
T × deprived neighborhood		0.193**		0.065		0.199**		0.058
and foreign background		(0.079)		(0.082)		(0.086)		(0.081)
T × woman		0.033		0.038		0.036		0.042
		(0.045)		(0.046)		(0.046)		(0.047)
Observations	1,260	1,260	1,260	1,260	1,260	1,260	1,260	1,260
R-squared	0.128	0.130	0.118	0.120	0.117	0.121	0.118	0.120

Source: survey of applicants. Linear probability model with sampling weights. No controls. No fixed effects. **Note** : Robust standard errors in parentheses. Significant differences at 10%, 5% and 1% are denoted by *, **, and ***.

Table 2.23: Robustness to different samples /specifications

	Interview				Recruitment			
	benchmark	unweighted	before-after	logit	benchmark	unweighted	before-after	logit
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	.131***	.117***	.125***	-1.885***	.016**	.028**	.032***	-4.127***
	(.035)	(.023)	(.027)	(.329)	(.007)	(.012)	(.010)	(.430)
Deprived neighborhood or foreign origin	-.028	-.027	-.022	-.301	.002	-.002	-.015	.115
	(.032)	(.026)	(.027)	(.328)	(.012)	(.014)	(.011)	(.537)
woman	-.022	.042	-.0002	-.234	.011	.015	.006	.521
	(.033)	(.028)	(.027)	(.344)	(.012)	(.014)	(.012)	(.540)
Anonymous resumes (T)	.042	.080**	-.007	.300	.025*	.017	-.0009	.823
	(.047)	(.038)	(.033)	(.412)	(.015)	(.018)	(.013)	(.569)
T x deprived neighborhood or foreign origin	-.100**	-.075*	-.041	-1.172***	-.035*	-.025	-.011	-1.194*
	(.044)	(.039)	(.034)	(.450)	(.020)	(.021)	(.015)	(.714)
T x woman	.028	-.055	.013	.298	.009	.008	.014	.108
	(.045)	(.040)	(.035)	(.474)	(.021)	(.022)	(.017)	(.745)

Source: survey of applicants. Linear probability model. No controls. No fixed effects. **Note** : Robust standard errors in parentheses. Significant differences at 10%, 5% and 1% are denoted by *, **, and ***.

Figure 2.1: Share of women among job-seekers, by position sought

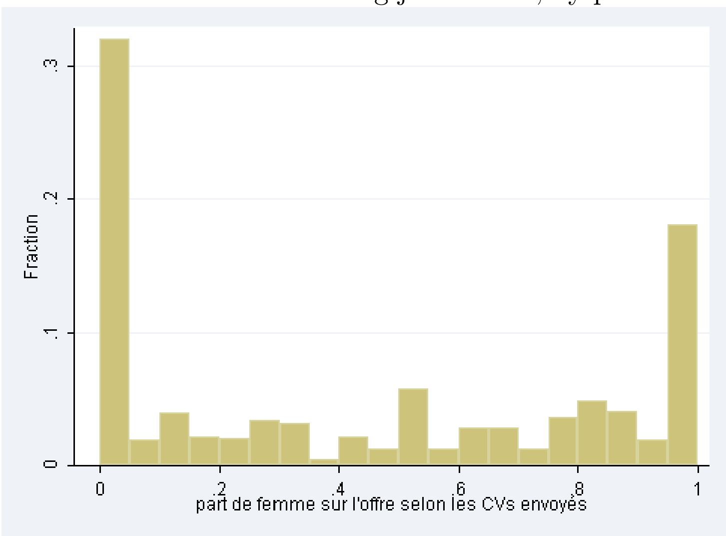
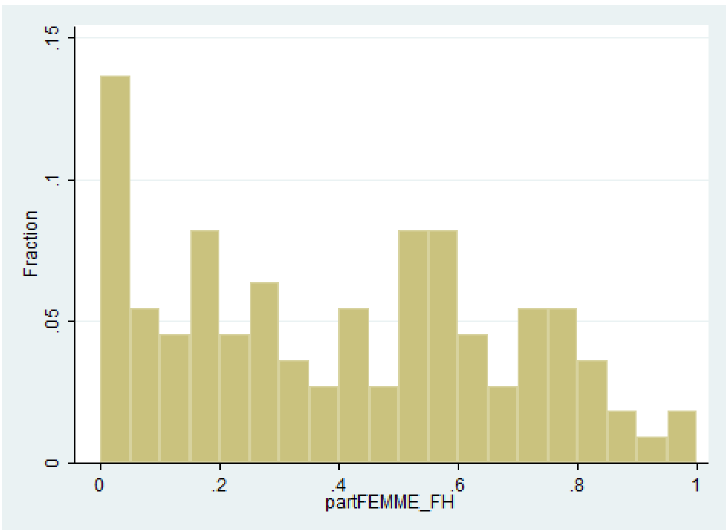


Figure 2.2: Share of women among resumes preselected by the PES, experimental offers



Chapter 3

The Effect of Potential Unemployment Benefits Duration on Unemployment Exits to Work and on Match Quality in France[1]

Contents

[1]This chapter is largely based on common work with Thomas Deroyon.

3.1 Introduction

There is a large body of empirical evidence on the impact of unemployment insurance (UI) generosity. Putting aside insurance provision, the empirical literature mostly focuses on impacts on labor market transitions from unemployment to employment. When unemployment benefits (UB) are more generous, reservation wages may increase and/or search effort may be lower. This leads to a decrease in unemployment exit rate to jobs. At the same time, unemployment benefits may affect the match quality, in a positive way as it encourages unemployed to wait for higher productivity jobs (see Marimon and Zilibotti (1997) and Acemoglu and Shimer (2000)) or in a negative way if human capital depreciates along the unemployment spell or employers discriminate candidates using unemployment duration. Effects on match quality are far less documented (see the review in Addison and Blackburn (2000)). Recent studies, such as van Ours and Vodopivec (2008), Card, Chetty, and Weber (2007), Lalive (2007), Centeno and Novo (2009) and Caliendo, Tatsiramos, and Uhlendorff (2009), do not find average effect on match quality. This paper provides evidence that effects on match quality are also limited in the French case. Compared to previous studies, this evidence is all the stronger that it concerns workers whose employability is particularly low (employed at most a year over the two years before unemployment registration). Those workers are likely to benefit the most from more generous UB. Low employability workers typically lack productive or job search skills that they could acquire thanks to extended Potential Benefit Duration (PBD). They are also likely to be financially constrained such that more generous UB would greatly change the value they attach to unemployment.

Our evidence is also all the stronger that we estimate the effect of a large increase in UB generosity. In a regression discontinuity design (RDD) similar to Card, Chetty, and Weber (2007), we estimate the impact of an increase from 7 to 15 months in potential benefit duration (PBD). In the French unemployment insurance system from 2000-2002, when workers are employed more than 8 months over the year before their job separation, they are entitled to 8 more months of UB: their PBD is more than doubled. This large increase makes our design very instructive: effects are expected to be large. However this large increase could also be a threat to the exogeneity assumption of our RDD. Indeed workers have huge incentives to cross the 8 month threshold. If they do so and actually accumulate just after the threshold, selection into treatment would be endogenous. Yet we do not find any mass point in the distribution of workers just after the threshold and we can be confident in the validity of our design.

Our result is robust to different measures of match quality: employment duration and hourly wage of the first job after unemployment exit. We complement those two standard indicators by the wage two years after unemployment entry. This enables us to compare short and extended PBD recipients at the same horizon, whatever the effect of PBD on unemployment duration. Medium term effects are relevant because accepting a stepping stone job could be in the end as efficient as longer job search for productive jobs.

Absence of match quality effect is all the more compelling that extension of PBD actually slows down unemployment exits to work. Unemployed with extended PBD wait longer before taking a job (roughly 2.5 months). Yet they do not find better jobs. The effect on unemployment exit to work starts early in the unemployment spell, even when both short and extended PBD unemployed receive benefits. This points to forward looking behaviors. However the effect is somehow stronger between 7 and 15 months after unemployment entry, when short potential benefits are expired but extended benefits are still paid. In addition, we verify that effects on registered unemployment duration are not only driven by claimants' obligation to register as unemployed, PBD extension also increases non employment duration. However this increase in non employment duration is half as large as the increase in registered unemployment duration and concentrated between 7 and 15 months after unemployment registration.

Our paper starts with a review of existing evaluation of UI generosity effect. Then we give background information on the institutional environment of job seekers on the French labor market. We present our data and describe our sample. In the next section, we motivate our regression discontinuity design. In the fifth part, we show that extended benefit duration tends to slow down unemployment exits. Finally we show that extended benefit duration do not have any strong effect on match quality.

3.2 Previous literature

Most of the empirical literature, including our paper, refers to the non stationary job search model described in van den Berg (1990). At every date of his/her unemployment spell, denoted t, the agent receives unemployment benefits $b(t)$ and draw with probability $\lambda(t)$ a wage offer from a distribution $F(w,t)$. He/she accepts the offer if the corresponding value of employment exceeds the continuation value of job search. This corresponds to a wage reservation strategy : unemployed accept the offer if the wage offered exceeds some value $\phi(t)$. van den Berg (1990) solves for the general strategy[2], but also considers the special case where the job offer probability

[2]The general reservation wage strategy is solution to the following differential equation with ρ the discount rate:

$$\phi' = \rho\phi - \rho b(t) - \lambda(t) \int_{\phi(t)}^{\infty} (w - \phi(t)) dF(w;t)$$

λ and the wage distribution $F(w)$ do not depend on time spent unemployed and where UB is a decreasing step function of time spent unemployed. The reservation wage is then a decreasing function of unemployment duration and the actual unemployment exit rate an increasing function of unemployment duration. In panel 3.1, we draw the stylized effects of extending the PBD from 7 to 8 months as predicted by the standard non stationary job search model. The left figure presents the effect on unemployment exit rate; on the right, the effect on reservation wage is drawn. Because longer PBD increases the value of unemployment, the reservation wage of extended PBD unemployed is always higher than that of short PBD unemployed. Of course this description of effects may be over simplistic, as the job offer arrival rate and the wage distribution may change over time because of skill depreciation or because of increasing search efficiency over time.

[INSERT FIG 3.1 HERE]

Empirical evidence on the negative effect of UI generosity on the unemployment exit rate is numerous. In his seminal work, Meyer (1990) identifies the effect of UI generosity in the US through variations across states. Since the adoption of more generous UI is potentially endogenous at the state level, Card and Levine (2000) propose to focus on exogenous variations in UI generosity due to targeted unanticipated policy change. Using the same identifying method, positive effects of potential benefit duration (PBD) on unemployment duration[3] are found in European countries, such as Germany (Hunt (1995)), Austria (Winter-Ebmer (1998), Lalive and Zweimuller (2004), Lalive, Ours, and ZweimÃ$\frac{1}{4}$ller (2006)), Poland (Puhani (2000b)), Slovenia (van Ours and Vodopivec (2006)), Finland (Kyyrä and Ollikainen (2008)), Portugal (Addison and Portugal (2008)) and France (Fremigacci (2010)). Other authors rely on discontinuities in the UI system to identify the effects. Those discontinuities are usually age thresholds, as in Lalive (2008), Caliendo, Tatsiramos, and Uhlendorff (2009) and Schmieder, von Wachter, and Bender (2012a). One exception is Card, Chetty, and Weber (2007) who use discontinuities based on past employment thresholds. We follow their strategy.

By contrast, empirical evidence on the effect of UI generosity on match quality is scarce and mixed (see the review in Addison and Blackburn (2000)). Using a structural model, Belzil (2001) finds that increasing the PBD by one week leads to an increase in subsequent employment duration by 0.5 to 0.8 days. Jurajda (2002) and Tatsiramos (2009) compare benefit recipients to ineligible unemployed and find large positive effects of eligibility on employment duration. Centeno (2004) estimates that a 10% increase in unemployment insurance generosity translates into a 3% increase in subsequent job tenure. In more recent studies, authors focused on identifying causal effects through difference in difference method (van Ours and

[3] Positive effects of replacement ratios are also found through difference in difference methods in Sweden (Carling, Holmlund, and Vejsiu (2001) or Bennmarker, Carling, and Holmlund (2007)) and in Finland (Uusitalo and Verho (2010)).

Vodopivec (2008)) or through regression discontinuities method (Card, Chetty, and Weber (2007), Lalive (2007), Centeno and Novo (2009), Caliendo, Tatsiramos, and Uhlendorff (2009), Schmieder, von Wachter, and Bender (2012b)). They do not find any average effects of PBD on subsequent wage, nor on employment duration. However, Centeno and Novo (2009) and Caliendo, Tatsiramos, and Uhlendorff (2009) show that match quality effects are heterogeneous. Centeno and Novo (2009) find that more constrained unemployed experience an increase by 3 to 8% in their earnings when PDB increases by 6 months. Caliendo, Tatsiramos, and Uhlendorff (2009) find that unemployed who find jobs just before their unemployment benefits run out accept less stable jobs than comparable unemployed who benefit from longer entitlement.

Our paper extends this empirical literature by estimating the effect of unemployment generosity on French workers with low employability[4]. As in Card, Chetty, and Weber (2007) and Centeno and Novo (2009), the effect is also estimated on younger unemployed than most existing regression discontinuities estimates which usually rely on age thresholds late in the worker's career. However, in contrast to Card, Chetty, and Weber (2007), we focus on workers with low employability. They have been employed at most one year over the two preceding year, where as, in Card, Chetty, and Weber (2007), workers have been employed around 2.5 years over the 5 year period before unemployment.

3.3 Institutional background

In France from 2000 to 2002, unemployed aged less than 50 years old could benefit from four different potential benefit durations (PBD). Entering into those different categories depends on past employment duration over a reference period. Job seekers with very long work experience could receive their unemployment benefits (UB) for up to 30 months, while the PBD is only 4 months for job seekers with the shortest employment duration (i.e. 4 months over the 18 last months). In our paper, we focus on the intermediate categories. Intermediate categories share the same reference period, one year before job separation, so that they can be easily compared in a regression discontinuity design. Job seekers, whose past employment duration is between 6 and 8 months, are entitled to 7 months of unemployment benefits; they will be referred to as short PBD job seekers. Job seekers whose past employment duration exceeds 8 months over the previous year will be referred to as extended PBD job seekers; they are entitled to 15 months of benefits.

[4] Dormont, Fougere, and Prieto (2001) study the introduction of decreasing replacement rate during the unemployment spell in France. Because the policy impacted all unemployed, they have not computed difference in difference estimates. Nor have they implemented regression discontinuity on the date of policy introduction. Our paper is thus to our knowledge the first causal evidence in the French case.

In all PBD categories, UB amounts are set according to the same rule. The replacement rate, i.e. the ratio of UB over former average wage, is degressive with the average wage on the year preceding the job loss. For average former wages around legal minimum wage, the replacement rate is around 66% gross, [5]. For average minimum wage twice as large as the minimum wage, the replacement ratio is 57.4% gross. Monthly UB amounts are capped at 5400 euros gross, one of the highest maximal amount in the OECD. The replacement rate rule changed in July 2001. Between January 2000 and June 2001, the replacement rate was smoothly decreasing with time elapsed in unemployment.[6] After July 2001, the replacement was constant during the whole benefit duration.

In the French UI system, senior job seekers (aged more than 50 years old) are given more generous UB and benefit from specific labor market programs policies (early retirement mechanisms, employers's subventions...). Job seekers working in temporary help agencies, or in particular occupations, such as technicians in the culture sector or artists, benefit also from specific UI rules. Even when those specific rules[7] are put aside, the French UI system is one of the most generous in the OECD (though less generous than the Danish and Dutch systems). The median maximal PBD among OECD countries is 12 months (24 months in France, see OECD summary table for 2005); the median replacement rate[8] in the OECD is 58% (67% in France); the median maximal monthly benefits payment is around 3 300 euros (it is 66% higher in France).

All UI recipients aged less than 50 years old have to register to the Employment Agency and abide by some rules to receive their benefits. They have to update their registration to the Employment Agency every month and, since July 2001, they have to meet a caseworker every six months. Monitoring is the same across PBD categories; thus the comparison between short and extended PBD cannot reflect differences in monitoring practices. Active labor market programs (ALMPs), such as counseling, training or skill assessment, are also available whatever the PBD category. ALMPs have become more frequent since July 2001 when semestrial meetings were introduced.

During their benefit claim, UI recipients are allowed to work in "side" jobs and add up the amounts earned with a fraction of their UI benefits. The spare UB can be received later on, such that the theoretical calendar date when benefits expire is delayed. Claimants working in "side" jobs remain registered at the Employment

[5]The monthly gross legal minimum wage was around 1 100 euros in the early 2000s.

[6]After 4 months, recipients were to lose around 15% of their benefits; after 10 months, there was a further 15 % decrease. Decreasing replacement rates make the difference in generosity between categories less important before than after July 2001. This difference turns out to be minor and does not induce great change in estimated effects.

[7]There are no specific rules for seasonal workers except that their replacement rate has been reduced for inflows after December 2002.

[8]For a single unemployed at the mean wage

Agency, indicating they are still looking for a better job. We thus consider them as still "unemployed" in the following. When UI benefits expire, unemployed may receive means-tested social assistance, called *Revenu Minimum d'Insertion, RMI*[9]. The amount depends on the family composition and earnings; in the early 2000s, a single adult could receive around 400 euros.

Aside from insurance rules, two features of the French labor market have to be highlighted. First, France ranks in the high-middle range among OECD indicator on employment protection strictness[10]. While strict, there is no discontinuity in employment protection around the 8 month threshold. There is no incentive for firms to separate from certain worker exactly before they reach 8 month of seniority. If it were the case, we could have concerns with the validity of our regression discontinuity design. Second the wage distribution is conditioned by a binding minimum wage. Around 10% of workers are paid at the minimum wage. Given that we focus on low qualified workers, the share of unemployed who face a rigid wage setting will be even higher and we expect match quality to be affected in its duration dimension.

3.4 The data

Our sample is selected from a French matched unemployment-employment registers data set (a complete description can be found in appendix 3.9). These data give information on the prior and posterior employment spells of benefit recipients, which is crucial to implementing our regression discontinuity design and to inspecting post unemployment job quality.

We select a flow of new unemployment benefit recipients who enter the Employment Agency from 2000 to 2002[11]. To avoid identification problems caused by the specific policies aimed at senior job seekers, we exclude from our sample people 50 years old and more (at date of registration). We also drop from our sample job seekers entitled to very specific insurance rules, such as recurrent workers from temporary help agency, artists, technicians working in the culture sector...

Short and extended PBD job seekers represent 28% of all new unemployment benefit recipients. The majority (63%) of recipients are entitled to 30 month benefits. Those recipients were employed during at least 14 months before registering as unemployed, i.e. a longer period than recipients in our sample of interest. Thus we

[9]Unemployed who worked 5 years over the last 10 years may receive *Allocation de solidarité Spécifique* which is less restrictive in terms of family earnings.

[10]See OECD employment outlook 2004

[11]We only retain new unemployment benefit recipients, meaning they do not have any residual benefits left from a former unemployment spell. Therefore, their potential benefit durations are directly linked with their employment spells since their last unemployment spell. When benefit recipients have residual benefits, a complicated rule extends their residual benefits according to their last employment spells.

will identify the impact of maximal benefit duration on recipients with relatively low employability. Table 3.8 in appendix 3.9 shows the job seekers' characteristics for different PBD categories. Job-seekers entering in intermediate categories (first column) are younger and have lower education and qualification than those entering *filière 5* (second column). The proportions of women and foreigners are higher. Their previous job position was less stable and less rewarding : only 14% had a permanent contract before their job separation, and their daily wage was 25% lower than in the broader group. Finally, they had spent almost a year unemployed during the last 3 years.

To implement our identification strategy, we need to observe past employment duration which conditions eligibility. This information is not precisely recorded in the unemployment registers. At the first interview, unemployed present some forms delivered by their former employers, job counselors verify their eligibility and usually record in the unemployment registers the minimal past employment duration of their corresponding PBD categories. We thus use employment registers to compute past employment duration. While better, information from the employment registers is not perfect: there is still some measurement error. First around one third of unemployed have no employment spells in the employment registers before unemployment. Second around 20% of unemployed have past employment duration as recorded in the employment registers which are not consistent with their potential benefit duration recorded in the unemployment registers. As displayed in table 3.8 in appendix 3.9, "consistent" job seekers have a stronger relationship to work than the unrestricted sample : they are more often qualified men, with high levels of education, higher former wages and longer past tenure, and they have been less often registered as unemployed in the past three years. This is no surprise as stable jobs are better reported in the employment registers. We also verified that unemployed persons looking for a job in agriculture or in care sectors are more likely to have inconsistent employment records. Their former employers, probably in the same sector, are not covered by the employment registers. Another issue related to measurement error could be that sample selection is different among unemployed with short or extended PBD, or more precisely that it is different locally around the threshold. This would bias our estimation. However the measurement error is symmetric around the threshold (as can be seen in figure 3.7 in appendix B). From now on, we exclude "inconsistent" workers from our sample[12]

Despite those inconsistencies, adding employment registers information to unemployment registers clearly increases the quality of measurement of unemployment registers exits to work. In our sample, 35% of the unemployed leave the Employment Agency reporting they have found a job. However, 29% of the unemployed leave the Employment Agency without reporting their new situation to their caseworkers, and the Employment Agency drops 9% of the unemployed for administrative

[12] An alternative strategy is to apply a fuzzy regression discontinuity design. Results are robust and reported in appendix B.

reasons (not showing up to interviews...). Those benefit recipients may have found a job. Indeed, 41% of the unemployed leaving the Employment Agency start a job recorded in the employment registers around their exit date[13]. Measuring exits destination in employment registers registers not only increases the levels of exits to jobs, it also affects their timing(a fact already highlighted by **?** and Boone and van Ours (2009)). It displaces usual exit spikes found before UB exhaustion to after exhaustion (see appendix 3.9). As expected, the raw comparison of both PBD group shows that the median registration duration is greater when PBD is longer (507 vs 306 days).

Adding employment registers also enables us to consider non employment duration, rather than registered unemployment duration. One advantage of non employment duration is that it does not depend directly on registration behavior which could be affected mechanically by claiming timing, especially around UB exhaustion date, or by PES monitoring rules. Actually, exhaustion spikes are smoothed when duration is measured as non employment (compare short PBD non employment exits in the right figure and register exits in the left one in panel 3.2). However one disadvantage is that non employment duration does not take into account that recently hired individual may still search for a job. In other words, it does not control for match quality. Then non employment duration may underestimate unemployment duration. This bias may be particularly important in the French context because unemployed have strong incentives to accept small jobs and stay registered at the PES (see previous section on institutional background). Unemployment register exits enable us to control for on-the-job search. This duration is central in our subsequent analysis, but we also test for robustness using non employment duration.

[INSERT FIG 3.2 HERE]

Finally, the unemployment registers do not contain any information about the exit jobs of benefit recipients. Employment registers help us to describe the employment duration of newly employed workers, their wages, and thus the part of their former wages they were able to recover.

In our sample, the median employment duration is 6 months[14]. The monthly job separation rate shows spikes at the usual temporary contracts durations : 6, 12 and 24 months (see the first graph in panel 3.3). Former job seekers with extended PBD stay longer in their new jobs than those with short PBD: the median of employment duration increases by 1 month between the two groups.

Half of job seekers gain more than 2 % of their former real hourly wages when they start a new job[15]. The wage gain is higher for the extended benefit duration job

[13]The corresponding employment spell should begin at most sixty days before or after the actual exit date and it should not end before it.

[14] Note that 14 % of new jobs spells are censored at the end of the data set (December 2004).

[15]Wage loss is computed as the ratio of starting wage over pre-unemployment wage as computed in the employment registers.

seekers (see the second graph in panel 3.3) : whereas more than one half of workers from the short benefit duration category do not recover their previous wage, the median wage gain is more than 3 % among extended duration job seekers.

[INSERT FIG 3.3 HERE]

The previous descriptive statistics show that job-seekers entitled to longer benefits take more time to find a new job (see panel 3.2). Their new jobs last longer and are more rewarding. Those differences shed some light on the link between unemployment generosity on the one hand and return to work and match quality on the other hand. However they could reflect the fact that recipients with extended PBD have worked during a longer period before becoming unemployed. They could have both observable and unobservable characteristics which make them less effective in job search, but more productive in their new jobs. In the following, we use a regression discontinuity design to address this potential endogeneity bias.

3.5 Identification strategy

Comparing individuals who have been randomly assigned extended potential benefit duration is the ideal design to estimate its causal effect. In a regression discontinuity framework (see Hahn, Todd, and Klaauw (2001)[16]), assignment to the extended benefit duration is locally random around the threshold of one forcing variable, here past employment duration. Then any difference in outcomes between recipients who are just below and just above the threshold can be attributed to the effect of extended potential benefit duration. The randomness assumption is impossible to test. However, there are ways to evaluate its credibility. First, precise manipulation of past employment duration is unlikely owing to the French institutional environment and the composition of our sample. Moreover, if there was precise manipulation of past employment duration, we should see some discontinuities in the distributions of the forcing variable and other covariates around the threshold. We present here the statistical tests we carried to assess this hypothesis.

3.5.1 Sample features that plea against precise manipulation

Local randomness of the forcing variable is not verified if some benefit recipients are able to precisely manipulate their employment duration. If that were the case, those individuals who manipulate employment duration would be just above the threshold, and the comparison of benefits recipients just below and just above the threshold would be biased. Actually, individuals who manipulate their employment duration are likely to have special characteristics highly correlated with unemployment exit rates, subsequent employment duration and wages.

[16] see for a practical guide Imbens and Lemieux (2008)

Manipulation could occur at different stages : at benefit registration, when employer and employee separate, or when they first meet. Our measure of past employment is robust to fraud at benefit registration. We observe past employment from an external source, not from administrative recordings at benefit registration, and we drop observations with inconsistent past employment history. Our sample excludes recurrent temporary workers and technicians working in the culture sector whose past employment certificates shown at benefit registration are more often erroneous than other unemployed groups (see the French *Cour des comptes* annual report 2010). Because most of the job seekers in our sample separate from temporary contracts, we believe that manipulation at job separation is less a concern than in the general case. The use of temporary contracts, and their extensions, is highly regulated in France. However we cannot exclude that employer and employee collude when they first meet, and set the contract duration so that it exactly extends the worker's past employment duration to meet the eligibility criteria to extended PBD. One argument which could limit the prevalence of collusion is that the employment prospects of our sample are structurally small. They are less educated and less qualified than the typical French worker. This should limit their ability to bargain.

3.5.2 Testing discontinuities in the forcing variable distribution

Turning to statistical argument, forcing variable manipulation can be checked by inspecting the population density around the eligibility threshold. If employment duration were precisely manipulated, recipients would accumulate just above 8 months. Figure 3.4 shows the forcing variable distribution around the threshold. The graph shows that there is no mass point just above the threshold.

[INSERT GRAPH 3.4 HERE]

To test formally for discontinuity in the population density (see McCrary (2008) for a reference), we estimate the following model :

$$N_d = \alpha + \delta I(d \geq \bar{d}) + (d - \bar{d}) \left(\delta_{-1} I(d < \bar{d}) + \delta_1 I(d \geq \bar{d})\right) + v \qquad (3.1)$$

where d is pre-unemployment employment duration[17], i.e. the forcing variable , N_d the population size of recipients with pre-unemployment employment duration d, $\bar{d}$ the threshold above which PBD is extended (8 months) and v is the error term. Thus $I(d \geq \bar{d})$ indicates whether individuals benefit from extended PBD. δ_{-1} and δ_1 capture linear dependencies between the forcing variable and the population size (allowed to be different below and above the threshold). Then the parameter δ captures the discontinuity in the population density at the threshold. We cannot reject the null hypothesis that δ is equal to 0 ($\hat{\delta} = 34$ with standard error 83). In

[17]In this regression, d is expressed in "weeks", more precisely in fourth of a month due to data limitation. In all subsequent regressions, d will be expressed in days.

this baseline estimation, the last point in figure 3.4 corresponding to jobs seekers who have worked one year before job separation has been discarded. The result of the test is robust to its inclusion. It is also robust to controlling for any "entire month" effect and for polynomials of past employment duration with higher degree (in the estimation above, the relation is assumed linear).

3.5.3 Testing discontinuities in covariates distributions around the threshold

Further evidence of the forcing variable exogeneity can be found by inspecting recipients' characteristics around the threshold. There indeed should be no discontinuities in the proportion of men, low qualified workers... around the threshold. Otherwise it would tend to prove that a certain part of the population manages to manipulate its past employment duration to gain longer benefits. To test for discontinuity we run several local linear regression discontinuity estimations on different windows around the threshold. The basic model we estimate has the same form as model 3.1 with the dependent variable being replaced by our characteristics of interest.

In table 3.1, the estimate of δ is reported for different populations around the threshold. In column 1, there are no restrictions on the sample included in the estimation. Unemployed in the extended PBD group may have worked one full year before claiming their benefits; they have thus quite different employment histories from the short PBD group who worked between 6 and 8 months over the preceding year. Indeed, when they contribute to the estimation, discontinuities are found in many characteristics : gender, age, education, marital status, qualification, unemployment history, contract type in previous work and quarter of separation. Those discontinuities highlight the fact that, in our design, regressions have to be local to be relevant [18]. Discontinuities persist when the estimation is restricted to a 4 month window around the threshold (column 2). But most of them vanish in the 2 month window estimation (column 3). In column 3, the estimation is restricted to unemployed who have worked between 7 and 9 months over the preceding year. It excludes individuals who worked exactly 6 months, a typical temporary contract duration (see the mass point in graph 3.3). The workers whose contract was exactly 6 months may be quite different from unemployed closer to the threshold and drive the discontinuities estimated on larger windows.

On windows smaller than 2 months, there are only two covariates discontinuities out of 21 independent covariates tested. We can thus be quite confident with our "no manipulation" assumption on those samples. Namely, we do not find strong evidence that there are more seasonal workers in the short PBD category[19]. It

[18] One other implication could have been that the linear assumption leads to misspecification. However taking that perspective is somehow less conservative (see results in tables 3.10 and 3.11 in appendix).

[19] This result may be related to the fact that workers in temporary help agencies have been dropped out from the sample.

would have been a serious concern as their unemployment exits are determined by calendar season and would bias our estimate of PBD extension. Although there are some differences in the seasonality of separation, they are not in the expected way: seasonal workers tend to separate after summer (??). Besides, there are no discontinuities in the share of unemployed looking for a temporary contract. This variable can be thought as a proxy for seasonal worker.

[INSERT TABLE 3.1 HERE]

3.6 The effect of potential benefit duration on unemployment exits to work

As we have already seen, there is an obvious relationship between employment history and the hazard rates out of unemployment. Our regression discontinuity design gives formal evidence of the causal impact of PBD on exits to work.

We start by illustrating the discontinuity we estimate in figure 3.5. To draw this graphics, we estimate the following Cox model of the hazard rate of unemployment exit to work, noted θ_t :

$$\theta_t = h_t \exp\left(\sum_{j \in 24..40} h_j I(d = j) + \gamma X\right) \tag{3.2}$$

where h_t is the baseline hazard rate (t is the time in weeks since the job seeker started claiming benefits), d is again the past employment duration (expressed in fourth of a month) and X represent a set of covariates. We graph in figure 3.5 the estimates of parameters h_j against past employment duration j on the sample restricted to the four month window around the threshold. There is a clear jump when crossing the threshold ($j = 32$). In the following, we estimate the size of the effect and its timing within the unemployment spell.

[INSERT GRAPH 3.5 HERE]

3.6.1 Estimating an overall effect of UI generosity on hazard rates

We first estimate the overall effect of PBD extension in the following regression discontinuity Cox model of the unemployment exit to employment :

$$\theta_t \quad = \quad h_t \exp\left(\delta I(d \geq \bar{d}) + (d - \bar{d})\left(\beta_{-1} I(d < \bar{d}) + \beta_1 I(d \geq \bar{d})\right) + \gamma X\right) \tag{3.3}$$

where d is the past employment duration expressed in days and all other variables have already been defined (in models 3.1 and 3.2). The model is estimated on the full sample and on subpopulations around the threshold. Given evidence on the tests on covariates, we prefer the estimations on the 2 or 1 month window. The estimates of parameter δ are reported in table 3.2. They are significantly different from 0 whatever the sample. The estimate on the narrowest window is significant at a higher level; it is somehow smaller than estimates on wider windows. The estimate on the 1 month window shows a decrease of the exit rate by 25% compared to 35% on the 2 month window.

[INSERT TABLE 3.2 HERE]

Table 3.12 in annex 3.11 shows some robustness checks. The results are robust, in strength and significance, when there are no covariate controls. Estimating the model with higher degree polynomials however alters the results. They are smaller in magnitude and no longer significant on the two month window.

One concern with the previous estimation could be that there are more seasonal workers in the short PDB category, so that the increase in exit rate of the short PDB category only reflects calendar effects. We have already partially addressed this issue as we do not find evidence of discontinuities around the threshold in proxies of the share of seasonal workers. Another way to isolate seasonal workers is to control for recalls, seasonal workers are typically recalled to their last employer. The last line of table 3.12 in annex 3.11 shows that the effect is robust when recalled unemployed are excluded from the sample.

In the previous model, the parameter δ captures the effect of benefiting from extended PBD rather than short benefits on the exits to job at any time in the unemployment spell. The model thus assumes that the effect does not depend on the timing of UI benefits, especially that the effect does not change close to benefit exhaustion, or when unemployment duration exceeds both short and extended benefits duration. However there are some evidence of spikes at benefit exhaustion (in figure 3.2), reflecting the fact that the finite duration of UI benefits makes job search non stationary. We next address this issue.

3.6.2 Estimating the effect of UI generosity on the dynamics of exits to job

During the first 7 months, both short and extended PBD benefits are paid. Extended PBD recipients are better off because they anticipate future benefits. Between both expiration dates (7 and 15 months), short PDB benefits are not paid any more[20]. After 15 months, both groups do not receive any UB. In the following

[20] Those theoretical expiration dates somehow minor the true expiration dates, as some unemployed may take up small jobs while registered as unemployed and delay the time when their benefits exhaust.

Cox model, the effect is allowed to vary along the unemployment spell:

$$\theta_t = h_t \exp\left(I(d \geq \bar{d})\left(\delta_0 I(t < t_0) + \delta_1 I(t_0 \leq t < t_1) + \delta_2 I(t_1 \leq t)\right)\right) \ldots \quad (3.4)$$

$$\exp\left((d - \bar{d})\left(\beta_{-1} I(d < \bar{d}) + \beta_1 I(d \geq \bar{d})\right) + \gamma X\right) \quad (3.5)$$

where all notations are already defined, except t_0 and t_1 which are the theoretical exhaustion dates of short and extended PBD (equal to 7 and 15 months). The estimations, again run on different windows around the threshold, are presented in table 3.3. We find that UI generosity has a negative effect on exits to jobs in the first 7 months, where all unemployed receive benefits (see line 1). This corroborates Card, Chetty, and Weber (2007) results that UI generosity has an effect on exits to jobs before the short PBD benefits exhaustion. However, we also find that UI generosity has a higher and robust effect on exits to job between 7 and 15 months when short benefits have expired and extended benefits have not. This "contemporaneous" effect is very strong in magnitude. In the 2 month window, it induces a 60% decrease in exit rate when PBD is extended. After 15 months, there is no significant difference between hazard rates to jobs of unemployed in the two categories, as all benefits have expired.

Table 3.13 in appendix 3.11 presents some robustness checks : we test whether the parameters values and significance are robust to the introduction of higher degree polynomials and to the exclusion of covariates. Effects after 7 months of unemployment are robust. Effects before 7 months are not robust to controls with higher degree polynomials.

[INSERT TABLE 3.3 HERE]

Although this estimation gives some insights on the dynamics of the effects, it is subject to dynamic selection bias[21] and should thus be interpreted with caution. The regression discontinuity design assures that, when entering unemployment, individuals just below and above the threshold are identical. However, if extended PBD affects differently two groups of unemployed, say that it reduces unemployment exits for group A, but not for group B, the composition of the unemployed population at any point later in the unemployment spell is different between short and extended PBD categories. The group A is over represented in the extended PBD category. When contrasting hazard rates, we mix two effects, one pure extended PBD effect and one composition effect.

3.6.3 Robustness : estimating the effect of UI generosity on non employment duration

We also estimate models 3.3 and 3.5 with non employment exit rate as the dependent hazard. Table 3.4 displays average effects in its upper part (model 3.3) and dynamic

[21] See for example Ridder and Vikstrom (2011)

effects in its lower part (model 3.5). Average effects on non employment exits are twice smaller than those on unemployment register exits rate (reported in table 3.2). The slow down in non employment exits is still significant at the 5% level on the 2 month estimation window. Effects on non employment exits are not significant before short PBD exhaustion and after long PBD exhaustion: they are concentrated between 8 and 16 months. This differs from dynamic effects on unemployment registers exits which appear before 7 month.

[INSERT TABLE 3.4 HERE]

All these elements tend to prove that UI generosity has a causal and negative impact on unemployment exits to employment. We next focus on estimation of PBD impact on match quality.

3.7 The Effect of potential benefit duration on match quality

We first consider match quality of the first job after leaving the unemployment registers. Match quality is captured by two components: hourly wage and employment duration. Wage is a classical proxy for match productivity, as it represents a fraction of the match surplus. However, as already mentioned, wage setting is quite rigid in France, so that the wage distribution is concentrated around the minimum wage. Thus we do not expect any strong wage effect of UI generosity. This first proxy can be fruitfully complemented by employment duration. Considering that employment is an experience good (Jovanovic), match quality is revealed as time goes by and signalled by continuing employment spell. We also consider hourly wage at a fixed horizon after unemployment registration (namely two years). This enables us to analyze longer term UI generosity effect, which abstract from the differences in the unemployment exit timing induced by the treatment. First we discuss bias arising because of selection in employment, and then present our results.

3.7.1 Selection into employment

To estimate the effects of PBD on match quality, we compare outcomes for job finders with short and extended PBD. This comparison may suffer from a well known bias due to different selection into employment across PBD categories (see Ham and LaLonde (1996)). Indeed the job seekers induced to exit unemployment because of shorter benefits duration may be a very special population with intrinsic characteristics that make them work in different jobs. Then comparing characteristics of jobs found after short and extended PBD unemployment spells results in comparing individual characteristics rather than measuring the causal impact of benefit length. Evidence in table 3.5 shows that this bias may not be so dramatic: the fraction of job seekers who find a job when leaving unemployment is the same across PBD categories[22].

3.7.2 Effects on the first job when leaving unemployment registers

In this section we restrict the sample to job seekers who find a job when leaving unemployment registers. The effect on starting wage is estimated using the following local linear regression discontinuity model :

$$Y = \alpha + \delta I(d \geq \bar{d}) + (d - \bar{d})\left(\delta_{-1} I(d < \bar{d}) + \delta_1 I(d \geq \bar{d})\right) + \gamma X + \epsilon \tag{3.6}$$

It has the same structure as the model 3.1. In addition we expand the set of covariates with respect to the previous analysis to account for labor market conditions at the time of unemployment exits (we include quarter dummies). Although those controls are potentially endogenous (unemployed with longer PBD may select into employment when labor market is tight), they control for the fact that, due to longer unemployment spells, labor market conditions are systematically different for unemployed in short and extended PBD. More precisely unemployed with extended PBD tend to exit later than short PBD (see graph 3.2). We also include dummies for calendar month of exit to control for seasonal labor market conditions. These dummies come on top of the seasonal worker dummy already included in previous analysis. Note that we also control for recalls to the same employer.

In model 3.6, we normalize starting wage by past employment wage. Our outcome of interest is thus the logarithm of the ratio between real hourly starting wage and real past employment wage. Differences highlighted by the descriptive statistics in figure 3.3 are not confirmed in the regression discontinuity estimation (line 1 in table 3.6). There are no significant effects of extended PBD, and parameter estimates are somehow volatile. Those results are robust when covariates controls are excluded, when polynomials of higher order are used, or when the wage is specified in level (see table 3.14 in appendix 3.11).

We divide employment spells after unemployment exits into two broad categories: those lasting strictly less than 8 months and those lasting more than 8 months. We thus distinguish between typical short temporary contracts and stable employment relations (panel 3.3). 8 months is an interesting threshold : it is indeed the extended PBD eligibility threshold. Then former job seekers who find a job lasting more than 8 months are entitled to extended PBD. Were they already in this category, 8 months can be understood as a renewal threshold. The effect on this binary variable is estimated using a linear regression discontinuity model[23] (equivalent to model 3.6).

[22]Even if the difference is not statistically significant, it could have been the case that it matters quantitatively. We have applied the bounding approach as in Lee (2009b). As expected estimates sets are quite large: 9 points on the wage equation estimated below.

[23]We can abstract from censoring issues : there are virtually no employment spells censored before 8 months.

In the 4 month window around the threshold (column 2 of the second line in table 3.6), extended benefit duration seems to have a positive effect on employment duration. Extending benefit duration increases the proportion of jobs lasting more than 8 months by 6 points. However, when the window is less than 2 months, the effect of extended benefit duration on employment duration is lower and not significant (columns 3 and 4). As a consequence, there is no evidence of a causal impact of extended PBD on employment duration. This conclusion is robust, when covariates are excluded (see table 3.15 in appendix 3.11) or when a Cox model of the hazard out of employment is estimated (see table 3.16 in appendix 3.11).

[INSERT TABLE 3.6 HERE]

3.7.3 Effects 2 years after unemployment entry

We now turn to medium term effects[24]. They are interesting for at least two reasons. First, as match quality is an experience good, it may be revealed by hourly wage progression as time goes by. Second, while unemployed with extended PBD may delay their unemployment exit, ex unemployed with short PBD may gain experience and move to other jobs, this process may also improve match quality. Medium term match quality is captured by hourly wage two years after unemployment entry. Again we do not consider as employed job seekers who work in side jobs but are still registered at the Employment Agency. As in the previous section, our analysis may suffer from a selection bias into employment. However we verify that the share of workers two years after registration is not affected by extended PBD (see line 2 in table 3.5). Results of the linear regression discontinuity model are presented in table 3.7. There are no significant discontinuities in the wage ratio due to PBD extension.

[INSERT TABLE 3.7 HERE]

3.8 Conclusion

In a regression discontinuity design inspired by Card, Chetty, and Weber (2007), we find that potential unemployment benefit duration has a significant and large impact on unemployment exits to work, but no impact on subsequent match quality. When job-seekers are entitled to 15 months of benefits instead of 7 months, only because they cross the 8 months past employment threshold, their exits to jobs are slowed down by around 27%, leading to an increase by two and a half months in unemployment duration. This effect is twice as large as RD estimate in Centeno and Novo (2009) where a 6 month increase in PBD induces job seekers aged 30 years old to stay unemployed around one month and a half longer. We find that

[24] We cannot consider long term effects due to data availability: the last cohort entering in our sample is observed during two years

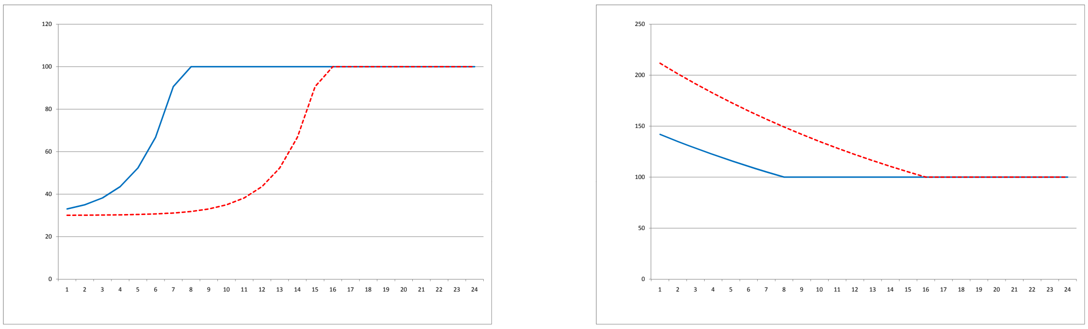
the unemployment exit slowdown is the strongest just after the short PBD expires. When both groups receive benefits, knowing that PDB would be extended induces a decrease by 22% in the exit rate, again this effect is twice as large as the RD estimate in Card, Chetty, and Weber (2007).

As in the recent literature using RD or DID estimates, we do not find any average improvement in subsequent match quality. Compared to the recent literature, our result can be seen as even stronger evidence of no match quality gain that our estimate of unemployment exits effects are twice as large as usual. Our result is all the stronger that our sample is made up of workers with low employability, who are not the typical population considered in most empirical literature.

Our result cannot be explained by the standard non stationary job search model with reservation wage unless skill depreciation or employer screening act as large counteracting forces to match quality gain.

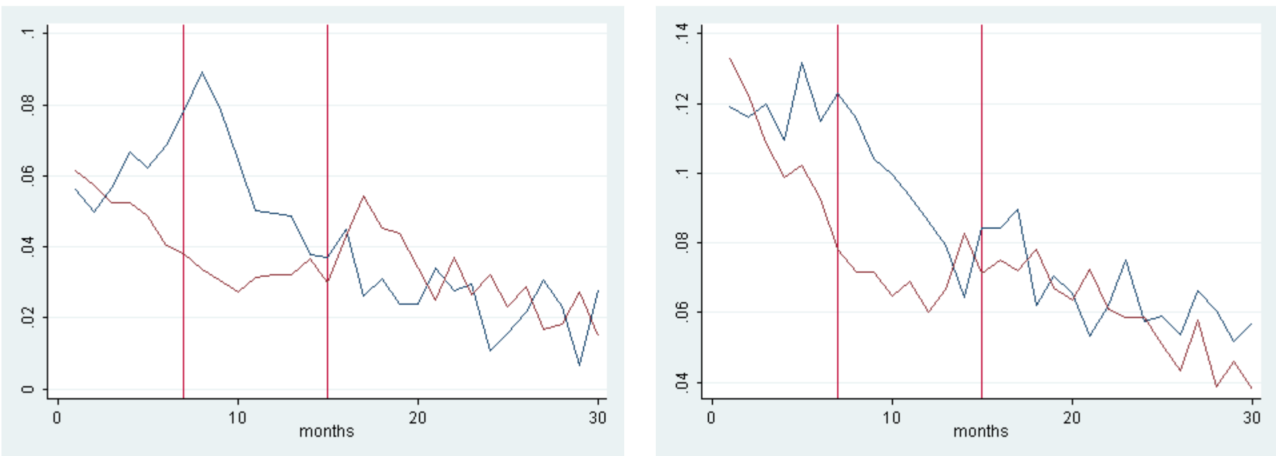
Graphics

Figure 3.1: *Stylized behaviors : Monthly unemployment register exit rates to jobs (on the left) and reservation wages (on the right)*



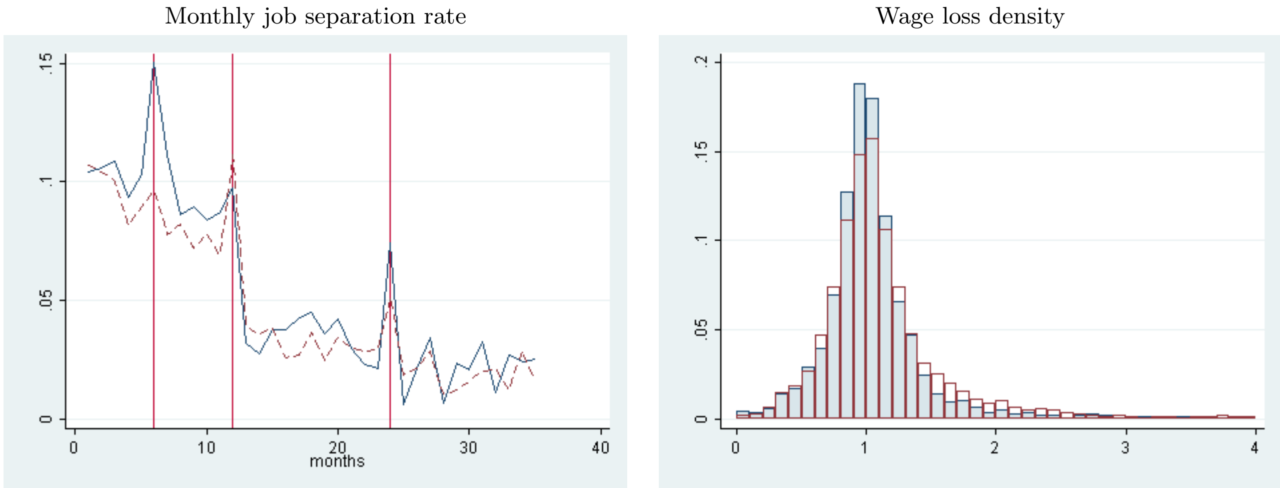
Reading: vertical lines represent dates of benefit exhaustion for short benefit duration (*filière 2*) and extended benefit duration (*filière 3*). The blue curve represents hazard rates for short benefit duration, the red curve for extended benefit duration

Figure 3.2: *Monthly unemployment register exit rates to jobs recorded in employment registers (on the left) and nonemployment exits (on the right)*



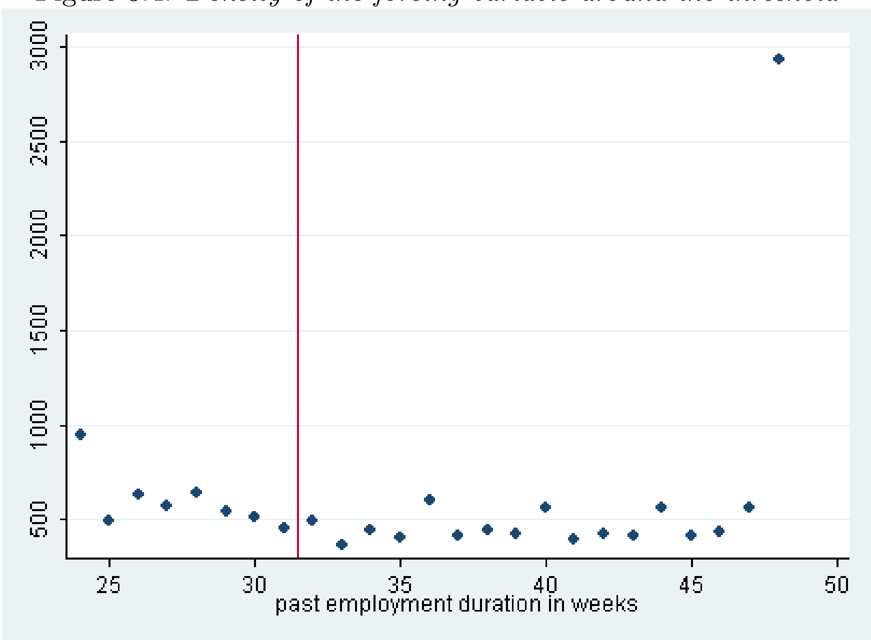
Reading: vertical lines represent dates of benefit exhaustion for short benefit duration (*filière 2*) and extended benefit duration (*filière 3*). The blue curve represents hazard rates for short benefit duration, the red curve for extended benefit duration

Figure 3.3: *Subsequent job quality*



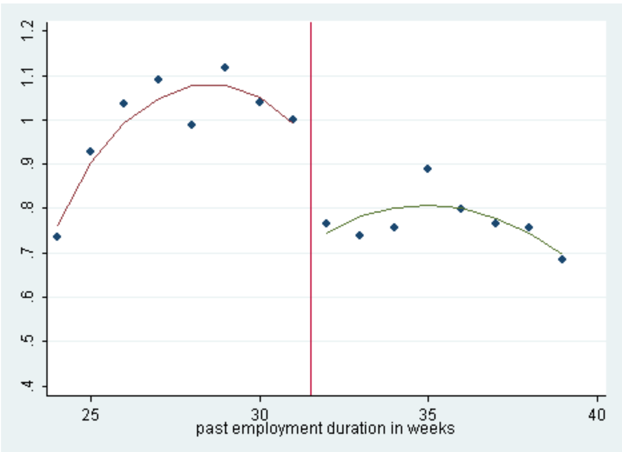
Reading: on the left graphics, vertical lines represent typical temporary contracts durations (6 months, 1 year, 2 years).

Figure 3.4: *Density of the forcing variable around the threshold*



Reading: on the x axis, we report past employment duration in "weeks" (more precisely in fourth of a month due to data limitation); it starts from 6 months (24 "weeks"), this is the minimum employment duration to enter *filière 2*. The vertical line represents the threshold between short and extended benefit duration. Mass points are found at typical contract duration (6 and 12 months).

Figure 3.5: *Unemployment exit to work hazard ratio and past employment duration in weeks*



Reading: on the x axis, we report past employment duration in "weeks"; it starts from 6 months (24 "weeks"), this is the minimum employment duration to enter *filière 2*. The vertical line represents the threshold between short and extended benefit duration. Here we compare unemployed in the 4 month window around the threshold).

Tables

Table 3.1: *Covariates discontinuity test on different windows around the threshold*

		Window around the threshold		
	All	4 months	2 months	1 month
	(1)	(2)	(3)	(4)
Man	-.074***	-.062***	-.054*	-.042
	(.018)	(.022)	(.032)	(.044)
Foreigner	.001	-.00006	-.019	-.013
	(.009)	(.011)	(.016)	(.022)
Age (log)	.031***	.020*	.009	-.0008
	(.009)	(.012)	(.017)	(.023)
Lower secondary education	.065***	.057***	.024	.039
	(.013)	(.016)	(.022)	(.030)
Vocational degree	.005	.016	.055*	.080*
	(.018)	(.022)	(.031)	(.043)
Upper secondary education	-.022	-.014	-.016	-.014
	(.015)	(.019)	(.026)	(.036)
Higher education	-.042***	-.058***	-.055**	-.088**
	(.015)	(.019)	(.027)	(.037)
Parent	.055***	.050**	.016	.021
	(.016)	(.020)	(.029)	(.039)
Married	.028*	.015	-.001	-.021
	(.016)	(.020)	(.029)	(.040)
Residence in great Paris region	-.022*	-.025	-.020	.016
	(.013)	(.016)	(.023)	(.030)
No qualification	.050***	.043**	.015	.033
	(.016)	(.020)	(.029)	(.039)
Low qualification	-.032*	-.004	.041	.018
	(.018)	(.023)	(.032)	(.044)
Intermediate profession	-.024**	-.027**	-.017	-.026
	(.010)	(.013)	(.018)	(.026)
Management	-.012	-.015*	-.016	-.016
	(.008)	(.009)	(.013)	(.018)
Previous hourly real wage	-.519	-.836	-1.070	-2.499
	(.955)	(.952)	(1.218)	(1.767)
Days unemployed during the last 3 years	48.513***	30.094**	19.879	12.293
	(11.166)	(13.817)	(19.748)	(27.572)
Previous work in service sector	.003	.013	.011	.012
	(.016)	(.020)	(.029)	(.040)
Looking for temporary contracts	-.013	-.016	-.012	-.020
	(.010)	(.012)	(.017)	(.024)
Previously on permanent contract	.025**	.010	.031	.048*
	(.012)	(.015)	(.020)	(.025)
Job separation during 1st quarter	-.015	-.033*	-.020	-.098***
	(.015)	(.019)	(.027)	(.037)
Job separation during 2nd quarter	-.024	-.006	-.048*	.003
	(.015)	(.018)	(.026)	(.035)
Job separation during 3rd quarter	.062***	.082***	.091***	.073*
	(.016)	(.020)	(.028)	(.038)
Job separation during 4th quarter	-.023	-.043**	-.024	.022
	(.017)	(.021)	(.029)	(.040)
Job separation before July 2001	.067***	.054**	.037	-.002
	(.018)	(.023)	(.032)	(.044)
Obs.	16692	8352	3837	1817

Local linear regressions. "Regression discontinuity" polynomials for the distance to the threshold of the forcing variable are of first order. Standard errors are robust to White heteroscedasticity.

Table 3.2: *Effect of extending potential benefit duration on unemployment exit rate to employment.*

	All	4 months	2 months	1 month
	(1)	(2)	(3)	(4)
Extending PBD	-.295***	-.283***	-.324***	-.225*
	(.053)	(.066)	(.094)	(.129)
Obs.	16692	8352	3837	1817

Cox model estimation. "Regression discontinuity" polynomials for the distance to the threshold of the forcing variable are of first order. All covariates tested in table 1 are included : gender, nationality, age, education, parental and marital status, residence in parisian region, qualification, past wage, separation reason, seasonal work dummy, past employment history, preferred sector, seasonality at registry and year dummies.

Table 3.3: *Dynamic effect of extending potential benefit duration on unemployment exit rate to employment.*

	All	4 months	2 months	1 month
	(1)	(2)	(3)	(4)
During the first 7 months	-.332***	-.228***	-.247**	-.184
	(.054)	(.070)	(.100)	(.139)
Between 8 and 15 months	-.956***	-.883***	-.947***	-.763***
	(.063)	(.084)	(.121)	(.169)
After 16 months	-.509***	-.053	.029	-.038
	(.077)	(.103)	(.156)	(.219)
Obs.	16692	8352	3837	1817

Cox model estimation. "Regression discontinuity" polynomials for the distance to the threshold of the forcing variable are of first order. All covariates tested in table 1 are included : gender, nationality, age, education, parental and marital status, residence in parisian region, qualification, past wage, separation reason, seasonal work dummy, past employment history, preferred sector, seasonality at registry and year dummies.

Table 3.4: *Effect of extending potential benefit duration on non employment exit rate.*

	All	4 months	2 months	1 month
	(1)	(2)	(3)	(4)
	Average effect			
Extending PBD	-.116***	-.167***	-.141**	-.137
	(.039)	(.049)	(.069)	(.095)
	Dynamic effect			
During the first 7 months	-.045	-.096*	-.048	-.062
	(.041)	(.052)	(.074)	(.101)
Between 8 and 15 months	-.380***	-.401***	-.481***	-.334**
	(.052)	(.066)	(.097)	(.138)
After 16 months	-.080	-.135*	-.062	-.235
	(.062)	(.076)	(.112)	(.160)
Obs.	16692	8352	3837	1817

Cox model estimation. "Regression discontinuity" polynomials for the distance to the threshold of the forcing variable are of first order. All covariates tested in table 1 are included : gender, nationality, age, education, parental and marital status, residence in parisian region, qualification, past wage, separation reason, seasonal work dummy, past employment history, preferred sector, seasonality at registry and year dummies.

Table 3.5: *Effect of extending potential benefit duration on survival in unemployment.*

	All	4 months	2 months	1 month
	(1)	(2)	(3)	(4)
Unconditionnal	-.008	.011	.017	.047
	(.018)	(.022)	(.031)	(.043)
2 years after registration	-.018	-.005	.015	.052
	(.016)	(.020)	(.029)	(.039)
Obs.	16692	8352	3837	1817

OLS estimation. Standard errors are robust to White heteroscedasticity. All covariates tested in table 1 are included : gender, nationality, age, education, parental and marital status, residence in parisian region, qualification, past wage, separation reason, seasonal work dummy, past employment history, preferred sector, seasonality at registry and year dummies.

Table 3.6: *Effect of extending potential benefit duration on match quality.*

	All	4 months	2 months	1 month
	(1)	(2)	(3)	(4)
Hourly wage ratio	..0009	-.009	-.030	.010
	(.024)	(.028)	(.040)	(.056)
Employment survival after 8 months	.034	.064**	.012	.020
	(.025)	(.032)	(.045)	(.062)
Obs.	7391	3797	1803	830

Standard errors are robust to White heteroskedasticity. "Regression discontinuity" polynomials in the distance between the threshold and the forcing variable are first order polynomials. All covariates tested in table 1 are included : gender, nationality, age, education, parental and marital status, residence in parisian region, qualification, past wage, separation reason, seasonal work dummy, past employment history, preferred sector, seasonality at registry and year dummies. Besides covariates capturing the seasonality and the business cycle at the exit date are included.

Table 3.7: *Effect of extending potential benefit duration on match quality 2 years after unemployment entry.*

	All	4 months	2 months	1 month
	(1)	(2)	(3)	(4)
Extended PBD	.007	.043	-.002	-.004
	(.029)	(.035)	(.051)	(.065)
Obs.	4546	2229	1058	489

Standard errors are robust to White heteroskedasticity. "Regression discontinuity" polynomials in the distance between the threshold and the forcing variable are first order polynomials. All covariates tested in table 1 are included : gender, nationality, age, education, parental and marital status, residence in parisian region, qualification, past wage, separation reason, seasonal work dummy, past employment history, preferred sector, seasonality at registry and year dummies.

3.9 Appendix A: Employment-unemployment registers

Our data set is based on the matching, at the individual level, of :

- the *Fichier Historique* (FH) of the French Public Employment Agency (ANPE), which records unemployment spells on a daily basis,

- and the employment registers or *Déclarations Administratives de Données Sociales* (DADS) built by the French Statistical Institute (Insee) from fiscal information collected annually at the establishment level. DADS records spells also on a "daily" basis.

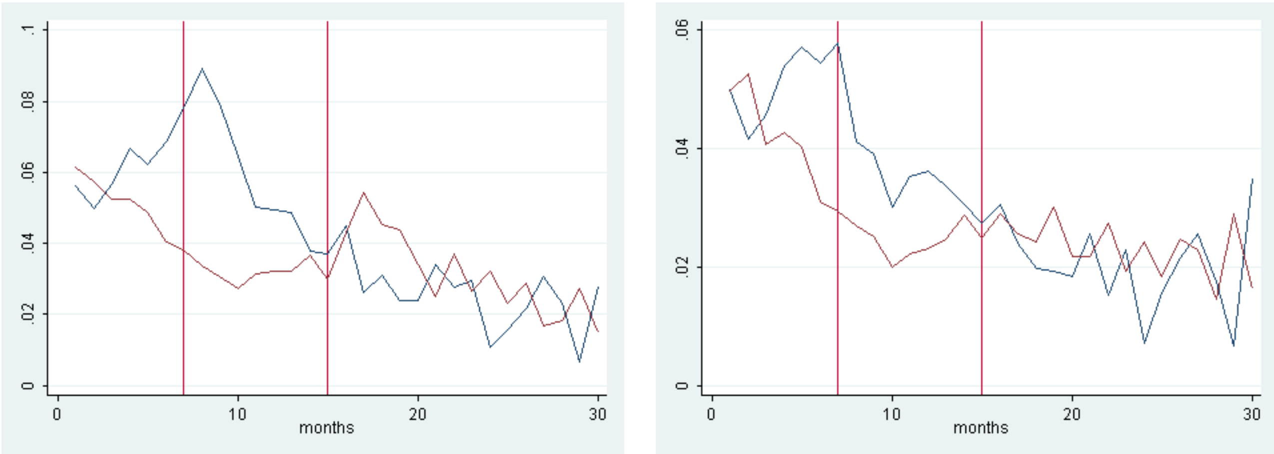
The employment registers cover around 85% of French wage earners. Civil servants from the French central adminstration (Ministers) and workers from the care sector or employed by a private person do not appear in the employment registers.
Due to legal restrictions (protection of private information), unemployed or workers have to satisfy two conditions to be included in the new data set:

- to be born in October of an even year,

- to be registered at least once in one or the other data set between 1999 and 2004.

For individuals in the matched sample, we observe all their unemployment and employment spells from January 1999 to December 2004. Spells are censored in December 2004. For individuals who appear at least once in the employment (resp. unemployment) registers between 1999 and 2004, employment (resp. unemployment) spells before 1999 are included (the employment registers start in 1976 and the unemployment registers in 1994).
All employment information (wage, duration and sector before or after unemployment) is taken from the employment registers except the type of previous contract (or separation reason) recorded in the unemployment register. Employment spells are within the same firm.
All unemployment information (namely duration, UB, desired type of contract) is taken from the unemployment registers. Socio demographics characteristics (gender, nationality, age, education, parental and marital status, place of residence, qualification) are recorded in the unemployment registers at the beginning of the unemployment spell.
We also consider non employment duration which is the time elapsed between two employment spells.
Despite some inconsistencies in the employment-unemployment history of individuals (discussed in the data section), the overall quality of the match is good. For example, the UB take-up rate, i.e. the fraction of claiming unemployed among eligible workers who separate from their employers, is similar when measured in the matched sample and in external sources.

Figure 3.6: *Monthly unemployment register exit rates to jobs recorded in employment registers (on the left) and to jobs reported in unemployment registers (on the right)*



Reading: vertical lines represent dates of benefit exhaustion for short benefit duration (*filière 2*) and extended benefit duration (*filière 3*). The blue curve represents hazard rates for short benefit duration, the red curve for extended benefit duration

Measuring jobs in employment registers not only increases the levels of exits to jobs (as explained in the data section), it also affects their timing (a fact already highlighted by **?** and Boone and van Ours (2009)). The lack of information due to missing job seekers' reports usually blurs the variations of exit rates to employment at benefit exhaustion and casts doubts on the existence of spikes at that time. The exit rate to jobs, as reported to the Employment Agency, does indeed rise and decline before benefit exhaustion (see the second graph in panel 3.6). However the exit rate to jobs, as recorded in the employment registers, rises before the end of benefit exhaustion and reaches a spike just after it (see the first graph in panel 3.6). This certainly highlights a change in the reporting behavior of job seekers at benefit exhaustion.

Table 3.8: *Effects of sample selection on covariates*

	Short or extended PBD *filières* 2 and 3	Long PBD *filière* 5	Final sample restricted *filières* 2 and 3
Man	0.46	0.49	0.48
Foreigner	0.09	0.06	0.07
Age (log)	29.58	32.28	28.75
Lower secondary education	0.21	0.14	0.15
Vocational degree	0.38	0.42	0.37
Upper secondary education	0.19	0.18	0.21
Higher education	0.19	0.24	0.24
Parent	0.34	0.43	0.28
Married	0.33	0.46	0.29
Residence in great Paris region	0.16	0.20	0.17
No qualification	0.31	0.20	0.27
Low qualification	0.47	0.50	0.49
Intermediate profession	0.07	0.10	0.09
Management	0.04	0.08	0.05
Previous hourly real wage	6.42	8.88	7.95
Days unemployed during the last 3 years	311.67	106.19	286.40
Attached to the service sector	0.71	0.72	0.72
Looking fir temporary contracts	0.08	0.08	0.08
Previously on permanent contract	0.14	0.40	0.15
Observations	31945	71184	16692

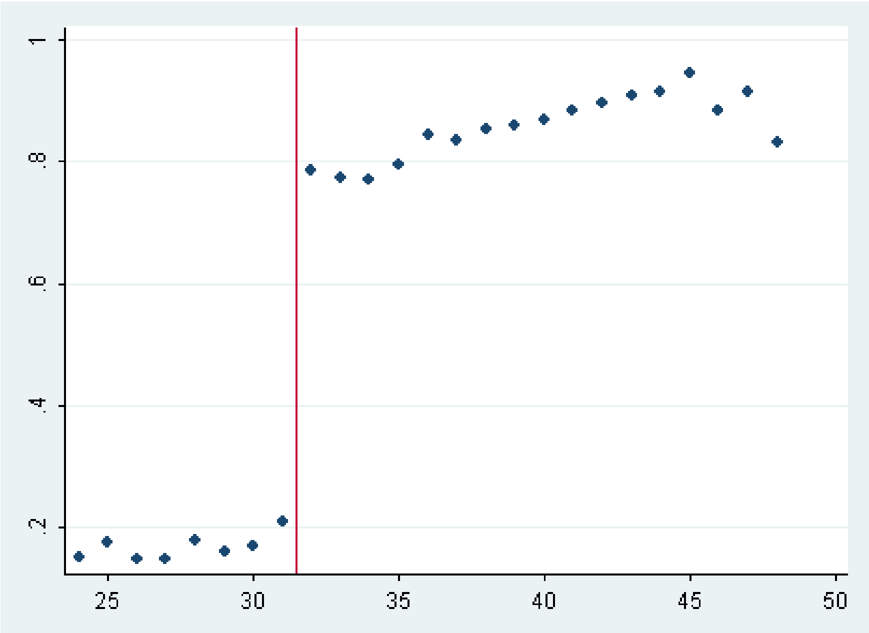
3.10 Appendix B: Fuzzy design

To test the influence of dropping ”inconsistent” workers, we estimate the effect of PBD using a fuzzy RD design on the whole population and compare the results with the effects of PDB estimated in a sharp RD design on the ”consistent” population. In the fuzzy regression design, the treatment is instrumented by the prediction using the forcing variable. We estimate effects on unemployment and non employment duration.

Graph 3.7 shows the evolution of the fraction of extended PBD recipients with past employment duration as recorded in the employment registers. As already mentioned, past employment duration is measured with error. Before the 7 month threshold, some unemployed are entitled to extended PBD, whereas they should not. After the threshold, some workers are entitled to short PBD, whereas they should benefit from extended PBD. Note that the error is symmetric.

Table 3.9 displays estimation results in the fuzzy (upper part) and sharp (lower part) design. Estimates in the fuzzy design are similar in magnitude to those in the sharp design. They are however less precisely estimated and significantly different from zero for both durations only in the 2 month window. Broadly speaking those results confirm that excluding ”inconsistent” workers does not bias severely our analysis.

Figure 3.7: *Actual extended PBD category.*



Reading: on the x axis, we report past employment duration in "weeks" (more precisely in fourth of a month due to data limitation); it starts from 6 months (24 "weeks"), this is the minimum employment duration to enter *filière 2*. The vertical line represents the threshold between short and extended benefit duration.

Table 3.9: *Fuzzy and sharp designs*

	All	4 months	2 months	1 month
	(1)	(2)	(3)	(4)
	Fuzzy design			
Unemployment	61.732***	75.584***	83.693***	50.947
	(13.616)	(18.189)	(25.493)	(37.246)
Non employment	23.722	57.640***	44.336	17.293
	(14.901)	(19.677)	(27.972)	(41.332)
Obs.	17794	10107	4774	2280
	Sharp design			
Unemployment	63.380***	76.821***	85.131***	80.454***
	(9.286)	(11.719)	(16.550)	(22.749)
Non employment	29.162***	52.934***	50.015***	50.233**
	(10.204)	(12.759)	(18.121)	(25.218)
Obs.	15039	8352	3837	1817

Standard errors are robust to White heteroskedasticity. All covariates tested in previous section are included : gender, nationality, age, education, parental and marital status, residence in parisian region, qualification, past wage, separation reason, seasonal work dummy, past employment history, preferred sector, seasonality at registry and year dummies.

3.11 Appendix C: Robustness

Table 3.10: *Covariates discontinuity test on different windows around the threshold (2nd order polynomials)*

	Window around the threshold			
	All	4 months	2 months	1 month
	(1)	(2)	(3)	(4)
Man	-.001	-.028	-.020	-.023
	(.027)	(.033)	(.046)	(.064)
Foreigner	-.016	-.018	-.020	-.010
	(.014)	(.017)	(.023)	(.033)
Age (log)	-.016	-.011	-.003	.005
	(.014)	(.017)	(.024)	(.033)
Lower secondary education	.0006	.011	.040	.029
	(.019)	(.023)	(.032)	(.045)
Professional degree	.007	.033	.058	.053
	(.026)	(.032)	(.045)	(.062)
Upper secondary education	.003	.013	-.028	-.022
	(.023)	(.027)	(.037)	(.052)
College education	-.011	-.051*	-.059	-.037
	(.023)	(.028)	(.039)	(.053)
Parent	-.006	-.009	.005	.062
	(.024)	(.030)	(.041)	(.056)
Married	.0004	-.028	-.024	-.008
	(.025)	(.030)	(.042)	(.058)
Residence in Parisian region	-.002	-.010	-.001	.028
	(.020)	(.024)	(.032)	(.042)
No qualification	-.011	.003	.016	.026
	(.024)	(.030)	(.041)	(.057)
Low qualification	.032	.049	.032	.020
	(.027)	(.033)	(.046)	(.064)
Intermediate profession	-.009	-.009	-.029	-.021
	(.016)	(.019)	(.027)	(.038)
Management	-.007	-.014	-.015	-.006
	(.012)	(.014)	(.019)	(.029)
Previous hourly real wage	-1.250	-1.165	-2.994*	-2.361
	(1.454)	(1.460)	(1.818)	(2.305)
Days unemployed during the last 3 years	-24.347	-5.971	1.492	12.926
	(17.019)	(20.640)	(28.742)	(41.015)
Previous work in service sector	.023	.007	.012	.155***
	(.025)	(.030)	(.042)	(.060)
Looking for temporary contracts	-.005	-.008	-.020	-.049
	(.015)	(.018)	(.025)	(.036)
Previously on permanent contract	.093***	.059***	.068***	.118***
	(.016)	(.021)	(.026)	(.030)
Job separation during 1st quarter	.015	-.006	-.088**	-.128**
	(.023)	(.028)	(.039)	(.055)
Job separation during 2nd quarter	-.047**	-.038	.002	.044
	(.022)	(.027)	(.037)	(.049)
Job separation during 3rd quarter	.064***	.081***	.089**	.174***
	(.023)	(.029)	(.040)	(.055)
Job separation during 4th quarter	-.032	-.037	-.003	-.090
	(.025)	(.030)	(.042)	(.059)
Job separation before July 2001	.047*	.016	.006	-.023
	(.027)	(.033)	(.046)	(.064)

Local linear regressions. "Regression discontinuity" polynomials for the distance to the threshold of the forcing variable are of 2nd order. Standard errors are robust to White heteroscedasticity.

Table 3.11: *Covariates discontinuity test on different windows around the threshold (3rd order polynomials)*

		Window around the threshold		
	All	4 months	2 months	1 month
	(1)	(2)	(3)	(4)
Man	-.034	-.042	-.044	-.025
	(.037)	(.043)	(.060)	(.088)
Foreigner	-.021	-.022	.006	.039
	(.019)	(.022)	(.030)	(.041)
Age (log)	-.002	.017	.013	.035
	(.019)	(.022)	(.031)	(.045)
Lower secondary education	.020	.037	.037	-.003
	(.026)	(.031)	(.043)	(.063)
Professional degree	.071**	.100**	.100*	.055
	(.036)	(.042)	(.059)	(.086)
Upper secondary education	-.015	-.049	-.010	.034
	(.031)	(.035)	(.049)	(.072)
College education	-.064**	-.070*	-.100**	-.044
	(.032)	(.036)	(.051)	(.073)
Parent	.036	.036	.087	.107
	(.033)	(.039)	(.053)	(.076)
Married	-.003	.0004	-.010	-.053
	(.034)	(.039)	(.055)	(.080)
Residence in Parisian region	-.013	-.017	.048	.070
	(.026)	(.030)	(.040)	(.053)
No qualification	.006	.028	.057	.007
	(.033)	(.039)	(.053)	(.078)
Low qualification	.054	.031	-.010	.075
	(.037)	(.043)	(.060)	(.087)
Intermediate profession	-.044**	-.034	-.027	.012
	(.022)	(.026)	(.036)	(.052)
Management	-.007	-.013	-.011	-.044
	(.016)	(.018)	(.027)	(.045)
Previous hourly wage	-2.407	-2.505	-2.740	.238
	(1.779)	(1.784)	(2.147)	(3.398)
Days unemployed during the last 3 years	-3.758	26.211	19.869	-6.780
	(23.353)	(27.172)	(38.497)	(57.408)
Previous work in service sector	.048	.013	.075	.138
	(.034)	(.040)	(.056)	(.084)
Looking for temporary contracts	-.008	-.029	-.044	-.032
	(.021)	(.023)	(.034)	(.054)
Previously on permanent contract	.080***	.038	.083***	.139***
	(.021)	(.025)	(.030)	(.037)
Job separation during 1st quarter	-.061*	-.081**	-.172***	-.144*
	(.032)	(.036)	(.052)	(.078)
Job separation during 2nd quarter	-.004	-.025	.062	.106*
	(.030)	(.035)	(.046)	(.063)
Job separation during 3rd quarter	.078**	.108***	.105**	.143*
	(.032)	(.038)	(.052)	(.077)
Job separation during 4th quarter	-.013	-.002	.005	-.105
	(.034)	(.040)	(.056)	(.082)
Job separation before July 2001	.030	.018	-.043	-.114
	(.037)	(.043)	(.060)	(.089)

Local linear regressions. ”Regression discontinuity” polynomials for the distance to the threshold of the forcing variable are of 3rd order. Standard errors are robust to White heteroscedasticity.

Table 3.12: *Effect of extending potential benefit duration on unemployment exit rate.*

	All	4 months	2 months	1 month
	(1)	(2)	(3)	(4)
without covariates	-.446***	-.383***	-.373***	-.246*
	(.052)	(.065)	(.092)	(.126)
2nd order polynomials	-.130*	-.246**	-.084	-.062
	(.079)	(.097)	(.132)	(.189)
3rd order polynomials	-.153	-.249**	-.127	-.178
	(.106)	(.126)	(.175)	(.257)
Excluding recalls	-.261***	-.250***	-.273***	-.178
	(.057)	(.071)	(.102)	(.141)
Obs.	16692	8352	3837	1817

Cox model estimation. All covariates tested in previous section are included (except in the first line): gender, nationality, age, education, parental and marital status, residence in parisian region, qualification, past wage, separation reason, seasonal work dummy, past employment history, preferred sector, seasonality at registry and year dummies.

Table 3.13: *Dynamic effect of extending potential benefit duration on unemployment exit rate.*

	All	4 months	2 months	1 month
	(1)	(2)	(3)	(4)
		without covariates		
During the first 7 months	-.472***	-.330***	-.290***	-.208
	(.053)	(.069)	(.099)	(.137)
Between 8 and 15 months	-1.084***	-.966***	-.980***	-.754***
	(.063)	(.083)	(.120)	(.167)
After 16 months	-.643***	-.119	-.036	-.057
	(.077)	(.103)	(.155)	(.216)
		2nd order polynomials		
During the first 7 months	-.403***	-.301***	-.086	-.142
	(.073)	(.098)	(.134)	(.189)
Between 8 and 15 months	-1.021***	-.952***	-.796***	-.723***
	(.078)	(.106)	(.147)	(.208)
After 16 months	-.568***	-.120	.172	-.007
	(.087)	(.119)	(.174)	(.243)
		3rd order polynomials		
During the first 7 months	-.688***	-.423***	-.253	-.400*
	(.089)	(.120)	(.167)	(.240)
Between 8 and 15 months	-1.288***	-1.068***	-.957***	-.966***
	(.091)	(.125)	(.176)	(.251)
After 16 months	-.815***	-.231*	.024	-.226
	(.098)	(.135)	(.196)	(.275)
Obs.	16692	8352	3837	1817

Cox model estimation. All covariates tested in previous section are included (except in the first line): gender, nationality, age, education, parental and marital status, residence in parisian region, qualification, past wage, separation reason, seasonal work dummy, past employment history, preferred sector, seasonality at registry and year dummies.

Table 3.14: *Effect of extending potential benefit duration on hourly wage .*

	All	4 months	2 months	1 month
	(1)	(2)	(3)	(4)
Without covariates	-.011	-.012	-.031	.028
	(.023)	(.028)	(.040)	(.056)
Wage in level	-.004	-.017	-.023	-.029
	(.015)	(.018)	(.025)	(.035)
2nd order polynomials -.034	-.023	.063	-.005	
	(.036)	(.041)	(.059)	(.083)
3rd order polynomials	.016	.032	.023	.022
	(.051)	(.055)	(.078)	(.111)
Obs.	7391	3797	1803	830

Standard errors are robust to White heteroskedasticity. All covariates tested in previous section are included (except in line 1): gender, nationality, age, education, parental and marital status, residence in parisian region, qualification, past wage, separation reason, seasonal work dummy, past employment history, preferred sector, seasonality at registry and year dummies. Besides covariates capturing the seasonality and the business cycle at the exit date are included.

Table 3.15: *Effect of extending potential benefit duration on employment survival at 8 months.*

	All	4 months	2 months	1 month
	(1)	(2)	(3)	(4)
without covariates	.043*	.073**	.019	.014
	(.026)	(.033)	(.045)	(.063)
2nd order polynomials	.017	-.006	-.016	.058
	(.038)	(.046)	(.064)	(.096)
3rd order polynomials	.048	.012	.062	-.106
	(.052)	(.060)	(.086)	(.137)
Obs.	7617	3913	1858	854

Standard errors are robust to White heteroskedasticity. All covariates tested in previous section are included (except in line 1): gender, nationality, age, education, parental and marital status, residence in parisian region, qualification, past wage, separation reason, seasonal work dummy, past employment history, preferred sector, seasonality at registry and year dummies. Besides covariates capturing the seasonality and the business cycle at the exit date are included.

Table 3.16: *Effect of extending potential benefit duration on employment duration.*

	All	4 months	2 months	1 month
	(1)	(2)	(3)	(4)
Baseline	-.032	-.047	.039	-.076
	(.057)	(.073)	(.101)	(.143)
without covariates	-.051	-.055	.032	-.082
	(.056)	(.070)	(.098)	(.135)
2nd order polynomials	-.022	.057	-.030	-.037
	(.085)	(.104)	(.144)	(.203)
3rd order polynomials	-.100	-.017	-.182	.457
	(.118)	(.137)	(.188)	(.293)
Obs.	6966	3563	1689	777

Cox model estimation. "Regression discontinuity" polynomials in the distance between the threshold and the forcing variable are first order polynomials. All covariates tested in previous section are included : gender, nationality, age, education, parental and marital status, residence in parisian region, qualification, past wage, separation reason, seasonal work dummy, past employment history, preferred sector, seasonality at registry and year dummies. Besides covariates capturing the seasonality and the business cycle at the exit date are included.

Table 3.17: *Effect of extending potential benefit duration on match quality 2 years after registry.*

	All	4 months	2 months	1 month
	(1)	(2)	(3)	(4)
without covariates	-.010	.032	-.019	.007
	(.029)	(.034)	(.050)	(.064)
2nd order polynomials	-.040	-.019	.050	-.108
	(.045)	(.052)	(.069)	(.080)
3rd order polynomials	.033	.037	-.043	-.154
	(.060)	(.066)	(.083)	(.124)
Obs.	4546	2229	1058	489

Standard errors are robust to White heteroskedasticity. "Regression discontinuity" polynomials in the distance between the threshold and the forcing variable are first order polynomials in the first line. All covariates tested in previous section are included : gender, nationality, age, education, parental and marital status, residence in parisian region, qualification, past wage, separation reason, seasonal work dummy, past employment history, preferred sector, seasonality at registry and year dummies.

Not for publication

We present in this technical appendix the simulation of the reservation wage strategy when the job arrival rate and the wage offer distribution are stationary. We assume that the wage distribution is uniform with upper and lower bound, $\bar{w}$ and $\underline{w}$ and that UB are set to $\bar{b}$ from date 0 to T and $\underline{b}$ from date T on. Then the differential equation can be written:

$$\phi' = \rho\phi - \rho b(t) - \lambda \int_{\phi(t)}^{\bar{w}} \frac{w - \phi(t)}{\bar{w} - \underline{w}} dw$$

$$\phi' = \rho\phi - \rho b(t) - \lambda \frac{(\bar{w} - \phi(t))^2}{2(\bar{w} - \underline{w})}$$

Since date T, the problem is stationary and $\phi(= \phi(T))$ is the constant solution $(< \bar{w})$ to the second order equation below :

$$\rho\phi - \rho\underline{b} - \lambda \frac{(\bar{w} - \phi)^2}{2(\bar{w} - \underline{w})} = 0$$

Let's call $\bar{\phi}$ and $\underline{\phi}$ the solutions of the RHS of the differential equation before T:

$$\rho\phi - \rho\bar{b} - \lambda \frac{(\bar{w} - \phi(t))^2}{2(\bar{w} - \underline{w})} = 0$$

Note that, assuming $\bar{b} > \underline{b}$, we have $\phi < \underline{\phi} < \bar{\phi}$. Then the differential equation before T can be written:

$$\frac{\phi'}{(\phi - \bar{\phi})(\phi - \overline{\phi})} = \frac{-\lambda}{2(\bar{w} - \underline{w})}$$

We can integrate this equality between date $t(> 0)$ and T.

$$\int_t^T \frac{1}{(\bar{\phi} - \underline{\phi})} \left(\frac{1}{(\phi - \bar{\phi})} - \frac{1}{(\phi - \underline{\phi})} \right) d\phi = \int_t^T \frac{-\lambda}{2(\bar{w} - \underline{w})} dt$$

$$\log\left(\frac{\phi(t) - \bar{\phi}}{\phi(t) - \underline{\phi}}\right) - \log\left(\frac{\phi(T) - \bar{\phi}}{\phi(T) - \underline{\phi}}\right) = \frac{-\lambda(\bar{\phi} - \underline{\phi})}{2(\bar{w} - \underline{w})}(t - T)$$

$$\frac{\phi(t) - \bar{\phi}}{\phi(t) - \underline{\phi}} = \frac{\phi(T) - \bar{\phi}}{\phi(T) - \underline{\phi}} \exp\left(\frac{-\lambda(\bar{\phi} - \underline{\phi})}{2(\bar{w} - \underline{w})}(t - T)\right)$$

$$\frac{\phi(t) - \bar{\phi}}{\phi(t) - \underline{\phi}} = \frac{\phi(T) - \bar{\phi}}{\phi(T) - \underline{\phi}} \exp\left(\frac{-\lambda(\bar{\phi} - \underline{\phi})}{2(\bar{w} - \underline{w})}(t - T)\right)$$

$$\phi(t) = \frac{\bar{\phi} - \underline{\phi}\frac{\phi(T) - \bar{\phi}}{\phi(T) - \underline{\phi}} \exp\left(\frac{-\lambda(\bar{\phi} - \underline{\phi})}{2(\bar{w} - \underline{w})}(t - T)\right)}{1 - \frac{\phi(T) - \bar{\phi}}{\phi(T) - \underline{\phi}} \exp\left(\frac{-\lambda(\bar{\phi} - \underline{\phi})}{2(\bar{w} - \underline{w})}(t - T)\right)}$$

$$\phi(t) = \frac{\bar{\phi} - \underline{\phi}}{1 - \frac{\phi(T) - \bar{\phi}}{\phi(T) - \underline{\phi}} \exp\left(\frac{-\lambda(\bar{\phi} - \underline{\phi})}{2(\bar{w} - \underline{w})} (t - T) \right)} + \underline{\phi}$$

Chapter 4

Non-response bias in treatment effect models[1]

Contents

4.1 Introduction

Sample attrition is a pervasive issue for surveys in social sciences. The damage appears particularly clearly in randomized trials: while random assignment to treatment creates a treatment group and a control group that are at the same time comparable and representative of the initial population, in the presence of sample attrition, however, the observed treatment and control groups may not be comparable anymore, threatening the internal validity of the experiment; and they may not be representative of the initial population, threatening its external validity. A variety of tools has been developed to correct for sample selection over the past

[1]This chapter is largely based on common work with Luc Behaghel, Bruno Crépon and Marc Gurgand.

decades, starting with seminal papers by Heckman (1976 and 1979) and turning less and less parametric up to the "worst-case", assumption-free approach developed by Horowitz and Manski (1995, 1998 and 2000).

Yet, the issue seems to have received somewhat less attention in applied work than endogenous selection into treatment[2]. This may have to do with the lack of so-called within-study comparisons to make the point that sample selection matters in practice, as LaLonde (1986) showed selection into treatment does; this may also have to do with the difficulty to provide credible solutions when facing sample attrition. The main purpose of this paper is to propose another approach to correct sample selection, at the crossroads of semi-parametric forms of the "Heckit" and of the bounding approach of Lee (2009). The main advantage of our approach is to yield point identification without requiring an instrument, but making the most of quite basic information on the number of attempts that were made to obtain response to the survey from each individual that responded. We compare the assumptions of our approach to those of the bounding literature (Horowitz and Manski (2000); Lee (2009b)), and stress the role of assumptions on the monotonicity of response behavior.

We then apply our sample selection correction in the context of a job search experiment, which can be viewed as a within-study comparison proving that sample selection can matter in practice. In the context of a job search experiment, exhaustive administrative is available, but a phone survey yields richer information – with low and unbalanced response rates. Using the administrative information, we show that selection into the phone survey is not as good as random, as it is correlated to potential outcomes. Moreover, point estimates suggest that the phone survey over-estimates the program's impact by about 50%. Applying the sample selection correction procedure closes most of the gap between the estimates in the full and in the selected samples. Bounds *à la* Horowitz and Manski (2000) or Lee (2009b) are, in this application, too wide to be very conclusive.

This paper contributes to a wide, 40-year old econometric literature on sample selection put in perspective by Manski (1989), and some 20 years later by Lee (2009b). Lee (2009b) illustrates how the two main approaches – the latent index selection model and the bounding approaches – can in some instances converge, when they use – more or less explicitly – the same monotonicity assumptions. An important distinction however, is whether the method implies using an instrument. The fact that an instrument is needed in order for the identification of the Heckit model not to depend on arbitrary parametric assumptions is often considered as a major drawback of the approach, as plausible instruments are rarely available.[3] Our

[2]The problem is often assumed away after showing that response rates do not differ too much across comparison groups.

[3]"Unfortunately, it is often very difficult to find such variables [instruments] in practice. In our wage example, theory would suggest that household variables like children and the income of the spouses are likely to influence the reservation wage, but unlikely to influence the gross offered wage and hence should [be used as instrument]. However, these household data are not always available, and even if they are, it is not guaranteed that these variables are good predictors of the

conclusion is that the instrument may not be the issue: perhaps counter-intuitively, actual information on response conditions, even if not random, are enough to identify the same parameter as with an ideal instrument that would randomly vary the data collection effort across individuals.

Our results are useful to better understand the value and limits of the latent threshold-crossing selection model. They should also be of interest to practitioners facing sample attrition. A recent paper by Kremer, Miguel, and Thornton (2009) shows severe differential attrition in the context of a randomized experiment. Using Lee (2009) bounds, they recognize that they unfortunately cannot conclude on the impact of the program for half of their sample. Our approach could in principle be applied in such cases, yielding more informative results under assumptions that are arguably not too far from Lee (2009b).

The next section presents our sample selection correction approach. Section 3 gives an application, and section 4 concludes.

4.2 Sample selection correction using number of calls

In this section, we develop a new approach to sample selection correction using standard survey information on the data collection process. We first introduce the framework, recall and extend existing results on selection correction with instruments. We then present our approach, discuss its assumptions and compare it to the bounding approach.

4.2.1 Framework and notations

We use the potential outcome framework. $Z \in \{0, 1\}$ denotes the random assignment into treatment and $y(Z)$ is the potential outcome under assignment (or treatment) Z.[4] The observed outcome is $y = y(1)Z + y(0)(1 - Z)$. The parameter of interest is the average treatment effect:

$$E(y(1) - y(0)) \tag{4.1}$$

If Z is independent from $(y(0), y(1))$, as can be assumed under random assignment, then the average treatment effect can be estimated by comparing the empirical counterparts of $E(y|Z = 1)$ and $E(y|Z = 0)$. Alternatively, it is obtained by OLS estimation of:

$$y = \beta_0 + \beta_1 Z + \epsilon, \tag{4.2}$$

where β_1 is implicitly defined by $E(\epsilon) = E(Z\epsilon) = 0$.

propensity to work (...). But even if they are, the household variables may well be associated with the offered wage (...)." (Puhani, 2000, p. 58). Or: "The practical limitation to relying on exclusion restrictions for the sample selection problem is that there may not exist credible 'instruments' that can be excluded from the outcome equation." (Lee, 2009, p. 1080).

[4]In this section, we assume perfect compliance: treatment is equal to assignment (equivalently, if there is imperfect compliance, we consider the intention-to-treat effect. Appendix 4.5.3 provides an extension of our approach to the case with imperfect compliance.

Attrition bias may arise from non observation of the value of y, resulting from non-response behavior (whether literally or as a result of missing observation in any sort of data). Define potential response under assignment Z as $R(Z) \in \{0,1\}$. Just as for the outcome, $R(0)$ represents response behavior when a person is untreated and $R(1)$ when she is treated. Observed response behavior is $R = R(1)Z + R(0)(1-Z)$.

When there is non-response, the observed mean value of y in treatment and control group measures $E(y|R=1, Z=1)$ and $E(y|R=0, Z=0)$ respectively. Therefore, the "naive" average treatment effect estimator (for instance the above OLS estimation on the respondant individuals) measures:

$$\begin{aligned} E(y|R=1,Z=1) - E(y|R=0,Z=0) &= E(y(1)|R(1)=1) - E(y(0)|R(0)=1) \\ &= E(y(1)-y(0)) + \Delta_1 + \Delta_2, \end{aligned}$$

where the first equality obtains if Z is independent from $(y(0), y(1))$ and from $(R(0), R(1))$, and:

$$\begin{aligned} \Delta_1 &= E(y(1)-y(0)|R(1)=1) - E(y(1)-y(0)) \\ \Delta_2 &= E(y(0)|R(1)=1) - E(y(0)|R(0)=1). \end{aligned}$$

The first source of bias, Δ_1, results from treatment effect heterogeneity. It is present whenever the average treatment effect on those who respond to the survey $(R(1)=1)$ is different from the effect in the whole population. The second source of bias, Δ_2, is a selection bias: it occurs whenever treated and control respondents are different in the sense that they have different counterfactuals $y(0)$. None of these terms can be directly estimated because they require $E(y(0)|R(1)=1)$ but $R(1)$ and $y(0)$ are not jointly observed.

Bias Δ_1 involves lack of *external validity.* Absent the selection bias, the "naive" estimator would be consistent for a population of respondents to a given survey, but may not extend to the general population. There is no way this bias can be avoided given only respondent outcomes are observed. In contrast, the problem raised by bias Δ_2 is one of *internal validity.* Even if our interest lies on $E(y(1)-y(0)|R(1)=1)$, this would not be estimated consistently if this second type of bias is present.

In the following, we restrict the parameter of interest to average treatment effect on respondents, and we consider hypothesis under which the selection bias, if present, can be corrected.

Given the fundamental identification issue that characterizes the selection bias ($R(1)$ and $y(0)$ are not jointly observed), point identification of the causal treatment effect requires restrictions. Following Heckman (1976, 1979), a standard approach to sample selection correction relies on the latent selection model, whose identification requires instruments, i.e. determinants of selection (here response behavior) that do not determine the counterfactual outcomes. We present a semi-parametric version of that model but argue that proper instruments are difficult to find. We will then show in the next section that, provided the survey records the number of calls after which individuals responded, point identification is obtained based on that same model, even in the absence of instruments in the standard sense.

We assume the following latent variable threshold-crossing selection model:

Assumption 1.

1. *(Latent variable threshold-crossing response model)*

$$R = \mathbf{1}(V < p(W, Z)), \tag{4.3}$$

2. *(Independence)*

$$W, Z \quad \perp \quad y(0), y(1), V. \tag{4.4}$$

3. *(Common support)* $p(W, 0)$ *and* $p(W, 1)$ *have non empty common support* $\boldsymbol{P}$ *as* W *varies. Denote* $\bar{p}$ *the upper bound of* $\boldsymbol{P}$*.*

Equation (4.3) adds some structure to the relation between response and treatment status, $R(Z)$. W is an instrument that drives response behavior and that is independent from treatment assignment and potential outcomes; as will prove useful in the next section, W can be thought as the maximum number of attempts of phone calls to survey one individual. p is an unknown function, and without any loss of generality, V follows a uniform distribution over $[0, 1]$, so that $p(W, Z)$ is the response rate as a function of W and Z. V is not observed and can be interpreted as the individual reluctance to respond to surveys. A good example for W would be random variations in the incentives given to answer.

Proposition 1. ***: Identification with an instrument for response behavior.*** *Under assumption 1,* $E(y(1) - y(0)|V < \bar{p})$ *is identified. then*

$$E(y(1) - y(0)|V \leq \bar{p}) = E(y|R = 1, W = w_1, Z = 1) - E(y|R = 1, W = w_0, Z = 0), \tag{4.5}$$

where w_0, w_1 *are characterized by* $p(w_0, 0) = p(w_1, 1) = \bar{p}$*,*

This proposition builds on well-known properties of index selection models and adapts the semi-parametric identification results of Das, Newey, and Vella (2003) to our setting with a binary regressor of interest and heterogeneous treatment effects. The proof is given in appendix 4.5.1.1. To interpret equation (4.5), it is useful to think of V as an individual reluctance to respond to surveys. As W and Z are assigned randomly, V is equally distributed across treatment and control groups, and across different values of the instrument, W. Given the response model in (4.3), the population of respondents is uniquely characterized by the value of $p(W, Z)$. If there are two couples $(w_0, 0)$ and $(w_1, 1)$ such that $p(w_0, 0) = p(w_1, 1) = \bar{p}$, then these two couples are two different ways to isolate the same subpopulation of respondents. Therefore, comparing y across these two subpopulations (treated and controls) directly yields the impact of Z (on this specific subpopulation); i.e. the average treatment effect for those job seekers with $V \leq \bar{p}$. The popular Heckman selection correction model is a version of this with several parametric restrictions.

Of course, equation (4.5) is only useful to the extent that there exists such an instrument W. Unless this is planned in advance, it is usually extremely difficult to extract from the data some variables that have credible exogeneity properties and significant power to influence response. Therefore, as suggested above, randomizing survey effort could be a natural way to generate instrument W. One could for instance randomly assign the number of attempts to call each individual[5], or the value of the gift promised to those who respond. However, randomizing data collection effort amounts to wasting part of the sample on which data collection effort is not maximal. In most context, survey cost is high and the number of observations is limited; this may explain why, to the best of our knowledge, survey effort is not randomized in practice.

4.2.2 Number of calls as a substitute to instruments

Equation (4.5) is only useful to the extent that there exists such an instrument W. Unless this is planned in advance, it is usually extremely difficult to extract from the data some variables that have credible exogeneity properties and significant power to influence response. Therefore, as suggested above, randomizing survey effort could be a natural way to generate instrument W.

In what follows, we consider one such particular instrument (still denoted W): the maximum number of calls allowed. The main insight of this paper is that, if this instrument is valid, then a useful strategy is to set it at the same value w_{max} for the whole sample, and to record the information on the number of calls after which each individual responded. This approach allows to identify the exact same quantities as in proposition 1, and increases statistical power compared to a case where W would have been randomly varied across individuals.

Proposition 1 shows that, to identify $E(y(1) - y(0)|V \leq \bar{p})$, we need to identify $E(y|R = 1, W = w, Z = z)$ and $\Pr(R = 1|W = w, Z = z)$ for $z = 0, 1$ and for different values of w (in the case of the maximum number of calls, W takes integer values: $w \in \{1, 2, ..., w_{max}\}$). We consider a standard survey where survey effort W is not varied: $W = w_{max}$ for all individuals (w_{max} is typically chosen so that the marginal benefit of allowing for one more call rate equals the marginal cost). However, $E(y|R = 1, W = w, Z = z)$ and $\Pr(R = 1|W = w, Z = z)$ are identified if the survey records N, the number of calls after which each individual responds.[6] **Indeed, it is *immediate* that the set of individuals i who would respond if they were called up to w times is identical to the set of people who, when called up to w_{max} times (with $w_{max} \geq w$), respond at or before the**

[5]Specifically, one could design the survey so that it is randomly decided that worker i will be called a maximum number of times W_i before considering him as a non respondent if he or she cannot be reached, and worker j will be called a maximum number of times W_j, etc.

[6]Specifically, $N_i = w$ for an individual i who did not respond at the first $w - 1$ phone calls, but who responded at the w-th call.

w**-th call.** This is true in the control and in the treatment group alike, so that :

$$\{i : R_i = 1|W_i = w, Z_i = z\} \quad = \quad \{i : R_i = 1|N_i \leq w, Z_i = z\}$$

This immediately implies

$$E(y|W \leq w, Z = z, R = 1) \quad = \quad E(y|N \leq w, Z = z, R = 1) \qquad (4.6)$$
$$\Pr(R = 1|W = w, Z = z) \quad = \quad \Pr(R = 1|N = w, Z = z), \qquad (4.7)$$

for $z = 0, 1$ and $w \in \{1, 2, ..., w_{max}\}$.

We can thus state our main proposition:

Proposition 2. ***: Identification with the actual number of calls leading to response.*** *Under assumption 1, completed by the definition of N just above, $E(y(1) - y(0)|V \leq \bar{p})$ is identified from the observation of y, Z and N:*

$$E(y(1) - y(0)|V \leq \bar{p}) = E(y|N \leq w_1, Z = 1) - E(y|N \leq w_0, Z = 0), \qquad (4.8)$$

with w_0, w_1 such that

$$\Pr(N \leq w_1|Z = 1) = \Pr(N \leq w_0|Z = 0) = \bar{p}.$$

Results useful for estimation and inference are in appendix 4.5.2. Equation 4.8 means that $E(y(1) - y(0)|V \leq \bar{p})$ is point identified by simply truncating the control or the treatment sample. Figure 4.1.a illustrates this process. Individuals are ranked according to their unobserved reluctance to respond, and treatment does not affect the ranking, so that below any level $\bar{V}$ of the latent reluctance to respond, people with the same characteristics are present in the control and the treatment group. Without actual instruments, the information provided by the number of calls before the person responded acts as a proxy for V and make it possible to identify the marginal respondents. They can therefore be removed from the treatment–control comparison, thus restoring point identification. Point identification is only "local", however, in the sense that it is only valid for a population of respondents (the respondents in the group with the lowest response rate or any subgroup who have responded after fewer phone calls).

Finally, the fact that the maximum number of calls is not actually varied (i.e., that the instrument W is set at its maximum value w_{max} for the whole sample) is not a problem; the information that would be generated by varying W is embedded in N. The distribution of N informs about the ranking of individual reluctance to respond. In turn, the advantage of setting $W = w_{max}$ is to increase statistical power: in that case, $E(y|N \leq w, Z = z, R = 1)$ is estimated for all $w \leq w_{max}$ on the full sample; if we were to subdivide the control and the treatment groups in w_{max} subsamples on which we would set $W = 1$, $W = 2$,..., $W = w_{max}$, the average sample size to estimate $E(y|R = 1, W = w, Z = z)$ (for all $w \leq w_{max}$) would be divided by the number of subsamples, w_{max}.

This may seem too good to be true, as it implies that standard surveys embed the information needed for semi-parametric correction of sample selectivity. To put this result in perspective, one has to discuss whether the maximum number of calls is indeed a valid instrument, which amounts to discussing whether assumption 1 is plausible in the context of that particular instrument:

- The latent index threshold-crossing model embeds implicit restrictions that are well understood in the literature (Vytlacil, 2002) $\Rightarrow$ we discuss their implications for the specific instrument W in subsection 4.2.3
- The common support assumption may seem problematic given the fact that W and N are discrete. However, there are tools in the literature (Lee, 2009) to identify sharp bounds in that case $\Rightarrow$ we apply them in subsection 4.2.4

We conclude this section by comparing our approach to existing approaches in the bounding literature.

4.2.3 Restrictions implied by the latent variable model

This may seem too good to be true, and there is indeed a caveat: the assumptions of the latent model are actually strong assumptions. This is best seen by applying Vytlacil (2002) equivalence results in order to translate the index model into the potential outcome framework. Vytlacil shows that the index model is equivalent to assuming the following monotonicity condition:[7]

for all w, w', z, z', either $R_i(w,z) \geq R_i(w',z')) \; \forall\, i$, or $R_i(w,z) \leq R_i(w',z')) \; \forall\, i$.

This condition is a form of monotonicity condition. It therefore warrants the usual criticism: assignment to treatment Z may encourage some individuals to respond to the survey, but discourage some others. But even if the impact of each instrument (W and Z) is monotonous, the Vitlacyl (2002) monotonicity condition is more demanding. In particular, it does not hold if some individuals are only sensitive to W, and some only to Z. Assume for instance that W takes only two values (each person is assigned to 1 or 2 attempts). There may be a person i_1 who responds to the first call anyway: $R_{i_1}(2,0) < R_{i_1}(1,1)$. By contrast, person i_2 is only available at the second call, but responds to the survey irrespective of treatment assignment: $R_{i_1}(2,0) > R_{i_1}(1,1)$. In that case, W and Z have monotonous impacts, but Vytlacil (2002) monotonicity condition is violated. Potential response behavior is not observed, so that the monotonicity condition cannot be directly tested. However, this counter-example shows that this is by no way an innocuous assumption. Overall, and not surprisingly, point wise non-parametric identification comes at the cost of additional assumptions. What is perhaps more surprising is that finding a plausible instrument is not really the issue: the issue lies in the response model itself – a model widely used, apparently fairly general, but clearly not assumption-free yet.

[7] See LATE monotonicity assumption L-2 in Vytlacil (2002). Note that Z in his notations denotes the instrument (possibly a vector); in our notations, W and Z are the two instruments.

4.2.4 Discreteness of the number of calls

Our main result, stated in proposition 2, holds under the common support assumption (when N can be thought as continuous). In practice, N is discrete. As a consequence, it is not always possible to identify the exact cut off and drop corresponding marginal respondents. Then one can bound the average treatment effect on the respondents in the least responding group (let's say the control group with response rate $\bar{p}$). The bounding approach we adopt takes advantage of the bounded support of the potential outcome, but Lee bounds could also be used[8]. In our case, y only takes value 0 or 1[9]. Even though the effect is no longer point identified, the bounds obtained are typically tighter than bounds previously discussed which do not use any information on the number of calls.

When N is discrete, we can find w_1 such that $\Pr(N \leq (w_1 - 1)|Z = 1) < \bar{p} < \Pr(N \leq (w_1)|Z = 1)$. Let us call $\Pr(N \leq (w_1)|Z = 1) = p^*$. Then proposition 2 can be restated as follows (proof can be found in appendix):

Proposition 3. ***: Identification with the actual number of calls leading to response being discrete.*** *$E(y(1) - y(0)|V \leq \bar{p})$ is set identified from the observation of y, Z and N with lower and upper bounds $(\underline{\Delta}(\bar{p}), \overline{\Delta}(\bar{p}))$ such that:*

$$\underline{\Delta}(\bar{p}) = \frac{p^* E(y|N \leq w_1, Z = 1) - (p^* - \bar{p})}{\bar{p}} - E(y|N \leq w_0, Z = 0) \quad (4.9)$$

$$\overline{\Delta}(\bar{p}) = \frac{p^* E(y|N \leq w_1, Z = 1)}{\bar{p}} - E(y|N \leq w_0, Z = 0) \quad (4.10)$$

$$(4.11)$$

4.2.5 Comparison with bounding approaches

It is useful to compare our approach to the alternative, increasingly influential approach to the sample selection problem: the construction of worst-case scenario bounds of the treatment effect. This comparison will shed light on the trade-off between releasing identifying assumptions and improving what can be identified.

The assumption-free approach proposed by Horowitz and Manski (1995, 1998 and 2000) requires both weaker hypotheses (response behavior does not need to be monotonic) and less information (the number of calls does not need to be observed).[10] It does however require the outcome of interest to be bounded; moreover,

[8] We prefer to present the bounded outcome support case, because outcome support is bounded in our application

[9] The following proposition can easily be extended to bounded support of the form $[\underline{y}, \overline{y}]$

[10] See in particular Horowitz and Manski (2000). Assume that y is bounded: $-\infty < y_{min} \leq y \leq y_{max} < infty$. In its simplest form, the approach is to consider two extreme cases. In the best case, the outcome of all non respondent from the control group is y_{min} and the outcome of all treated non respondents is y_{max}; vice-versa in the worst case. If non respondents are in proportion nr_0 (resp. nr_1) in the control (resp. treatment) group, then the width of the identified interval is $(nr_0 + nr_1)(y_{max} - y_{min})$.

as illustrated by Lee (2009) or in the application below, it may generate very large bounds if response rates are not very high.

The approach suggested by Lee (2009) is much closer to our approach. It does not require y to be bounded. It provides tight bounds on treatment effects under the assumption that selection into the sample is monotonic, i.e., considering response $R(Z)$ as a function of assignment to treatment, $R(1) \geq R(0)$ for all individuals. The bounds are given by proposition 1a in Lee (2009, p. 1083). The width of the identified set can be substantially smaller than in Horowitz and Manski (2000), as it depends on the difference in response rates between the control and the treatment group, rather than their sum. Point identification is achieved when response rates are balanced.

Let us first compare Lee (2009) assumptions with our assumptions above: (i) our approach requires observing the actual survey effort leading to response; (ii) the two approaches require independence of assignment with regard to potential outcomes; (iii) both approaches impose monotonicity conditions on potential response behavior, but in our approach, the monotonicity condition is stronger as it bears jointly on the impact of assignment to treatment and on the impact of survey effort. Note that none of the two approaches requires y to be bounded. To sum, the two sets of assumptions are close; the main difference is that our assumption 1 implies a stronger form of monotonicity on response behavior.

Concerning identification results, the two approaches lead to point identification when response rates are balanced. Actually, when response rates are balanced between treated and controls, Lee's monotonicity assumption implies that respondents in the two groups represent the exact same population: there is no sample selection issue to start with. When response rates are not balanced, Lee (2009) yields set identification, while our approach still allows point identification. Figure 4.1.a (commented above) and 4.1.b illustrate the difference. In the two cases, individuals are ranked according to their unobserved propensity to respond, and treatment does not affect the ranking, so that at a given level V corresponding to a given response rate, individuals in the control and the treatment groups are comparable. Without instrument and without additional assumptions, Lee's approach does not allow to identify who the marginal respondents are, so that their outcome has to be bounded, yielding set identification. By contrast, as discussed already, the information provided by the number of calls before the worker responded acts as a proxy for V and makes it possible to identify the marginal respondents. Removing them from the treatment - control comparison restores point identification.

4.3 Application

In this section, we analyze non response in the context of a job search experiment (which actually initiated the research in this paper). We briefly present the program that is evaluated, the data, and evidence on sample selection bias. We then implement the sample selection correction proposed in this paper, and compare it to alternative (bounding) approaches. Our correction appears to reduce the sample selection bias, whereas the identified intervals from the bounding approaches are too wide to be conclusive here.

4.3.1 The program and the data

The phone survey used in this application took place in the context of a job search experiment (presented in more details in Behaghel, Crepon, and Gurgand (2012)). In 2007-08, the French unemployment benefit provider (Unédic) mandated private companies to provide intensive counseling to job seekers. To be eligible, job seekers had to be entitled with unemployment benefits for at least 365 days. The program was implemented as a randomized control trial: eligible job seekers were randomly assigned to the standard track (control group, with less intensive counseling) or to the treatment.[11]

Administrative data (i.e. the unemployment registers) are a first source of information to measure the program impact. We use one key variable, *exit*, which indicates whether the job seekers exited registered unemployment before the phone survey (before May 2008). Because individuals are benefit recipients we can be quite confident that an exit from the registers, which imply the suspension of benefits, is meaningful and related to a transition to employment (this view can be challenged when unemployed are not eligible to benefits; see Card, Chetty, and Weber (2008)). However, the administrative data does not allow to measure other relevant dimensions of impact, such as job quality on which the program put strong emphasis. To measure these dimensions, a phone survey was run in March and April 2008. The initial sample included around 800 job seekers out of the 4,300 individuals who had entered the experiment between April and June 2007 (see table 4.1). Job seekers had therefore been assigned for about 10 months when they were surveyed. The sample was stratified according to the job seekers' random assignment and to whether they had signed or not for an intensive scheme.[12] The interviews were conducted by an independent pollster mandated by the French Ministry of Labor. The questionnaire was rather long (a maximum of 103 questions, for an estimated average time of 20 minutes). Detailed questions were asked upon the track followed when unemployed (what they were proposed, whether they accepted or not, why, what they did,...) and on the current employment situation.

The response rate to the phone survey is low: Out of 798 individuals initially in the sample of the survey, only 57% actually answered to the poll (see Table 4.1). This motivates investigating the risk of sample selection bias. To do so, we use the exhaustive administrative data as a benchmark: are the results on *exit* the same if one considers the full sample as if one restricts the analysis to the sample

[11]Participation to the intensive scheme was not compulsory, so that compliance was imperfect. For the sake of simplicity, we focus on the intention-to-treat effect in this section, and refer to the appendix for a generalization to the identification and IV estimation of local average treatment effects.

[12]The analysis uses survey weights accordingly.

of respondents to the phone survey? Table 4.2 shows OLS estimates of intention-to-treat effects in the two samples. Estimated effects are about 50% larger when the sample is restricted to the phone survey respondents (a 13.6 percentage point impact, compared to 9.6 percentage points in the full sample). This is suggestive of a quantitatively significantly bias. Note however that the difference could come either from treatment effect heterogeneity (component Δ_1 in equation 4.3) or from sample selection bias (component Δ_2). A caveat is that standard errors are large, so that the difference is not statistically significant and could simply be due to sampling variations.

Table 4.3 provides evidence that sampling variations are not the only reason why the full sample and the sample of respondents yield different results: it shows that response behavior is statistically correlated with exit from unemployment registers. For instance, in the control group (line 1), there are sizable differences in exit rates between those who respond (column 2) and those who do not respond to the long telephone survey (column 1). The exit rate is 14.5 percentage points lower among respondents than among non respondents (column 4). As a consequence, considering the respondents only to estimate the exit on the whole population leads to a 7.2 percentage points (about 15%) downward bias (column 5). The fact that the phone survey over-represents job seekers with lower employment prospects can be interpreted in various ways: for instance, job seekers who have found a job are harder to reach, or they do not feel they have to respond to surveys related to the public employment service anymore. There is similar evidence of a downward bias for job seekers assigned to treatment.

Another important piece of evidence that non response is not "as good as random" in the phone survey is given by table 4.4, which shows that non response is correlated with treatment assignment: job seekers respond more when they are assigned to the private scheme than when assigned to the control group (the response rate increases by 13.7 percentage points, with a standard error of 4.0).

To sum up, the low and unbalanced response rates cast doubts on the validity of the phone survey in order to measure the program impact. These doubts are reinforced by the comparison with the exhaustive administrative data. In what follows, we implement different approaches to control for attrition bias in this data, and check whether these corrections close the gap between results with the full sample and results with respondents to the phone survey only. Recall that due to treatment effect heterogeneity it could be the case that the true effect on respondents (estimated with sample selection correction) actually widens this gap. Comparing estimates on the whole population and our corrected estimates cannot be a formal test for validity of our correction.

4.3.2 Selection correction

Table 4.5 displays estimates based on three correction approaches. In the first two columns, we recall the estimate on the whole population and the estimate on respondents. In columns 3 and 4, we report "bounding" estimates. The Horowitz and

Manski (2000) bounds and the Lee (2009) bounds are large, so that the telephone survey brings limited information.

In the last columns, we report estimates derived from our proposed correction method[13]. In columns 5 and 6, we abstract from the discreteness of the number of calls and present point estimates under two alternatives for the trimming threshold. In column 7, we take into account the discreteness and present bounds. In order to implement the correction, we need to find the number of calls at which to truncate the treatment group and restore the balance with the control group. Figure 4.2 displays response rates according to the number of phone calls and the assignment status. To restore the initial balance between experimental groups, the sample needs to be truncated between 6 and 7 phone calls in the group assigned to treatment. The corresponding point estimates based on proposition 2 are compatible with the whole population. They tend to close the gap between average treatment effect on respondents and on the whole population. Taken at face value, this implies that treatment effect heterogeneity is less an issue than internal validity of the effect on respondents. Standard errors in columns 5 and 6 are computed taking into account uncertainty due to the trimming procedure; hence they are larger than usual standard errors as in columns 1 or 2.

In column 7, we actually cut the sample between 6 and 7 phone calls in the treatment group. We assume two polar situations for marginal workers who respond after 7 phone calls when treated and would not have responded had they be in the control group. If we assume that they are employed (unemployed), we obtain a lower (upper) bound of the treatment effect. Even though the resulting identified set is quite large (2.6 points), it turns out to be strictly between the effect on the whole population and the naive effect on the respondents. More interestingly, the identified set is far smaller than usual bounding estimates (around 20 points).

4.4 Conclusion

In this paper, we argue against the view that finding plausible instruments is the key impediment to sample selection correction models in the line of the Heckman (1976, 1979) model. If that model is correct, basic information on the number of calls (or number of visits) that were performed before the individual responded is enough to obtain point identification of treatment effect, even in a semi-parametric model with heterogeneous treatment effect and a flexible specification of the latent threshold-crossing selection equation. The somewhat counter-intuitive result is that, despite the fact that reluctance to respond may well be correlated with potential outcomes, the actual effort made to get a response contains the same information as if survey effort was randomly allocated to individuals.

If the instrument is not the issue, it does not mean that there is no issue with such sample selection correction models. The true cost, however, lies in the restrictions that the model implies on response behavior. Clearly, if bounding approaches

[13] Estimates and standard errors are derived in appendix.

yield sufficiently narrow identified sets, they should be preferred as they imply less stringent restrictions. However, Horowitz and Manski (2000) bounds are quite large when response rates are below 80%, which is by no way the exception in social sciences. And the assumptions made by Lee (2009) are not so different from ours: extending the monotonicity assumptions may not be such a large cost compared to the substantial gains in terms of identification in cases where response rates are unbalanced, as in our application or in Kremer, Miguel and Thornton (2009).

Figure 4.1: Identification under Lee (2009) bounds and proposition 2

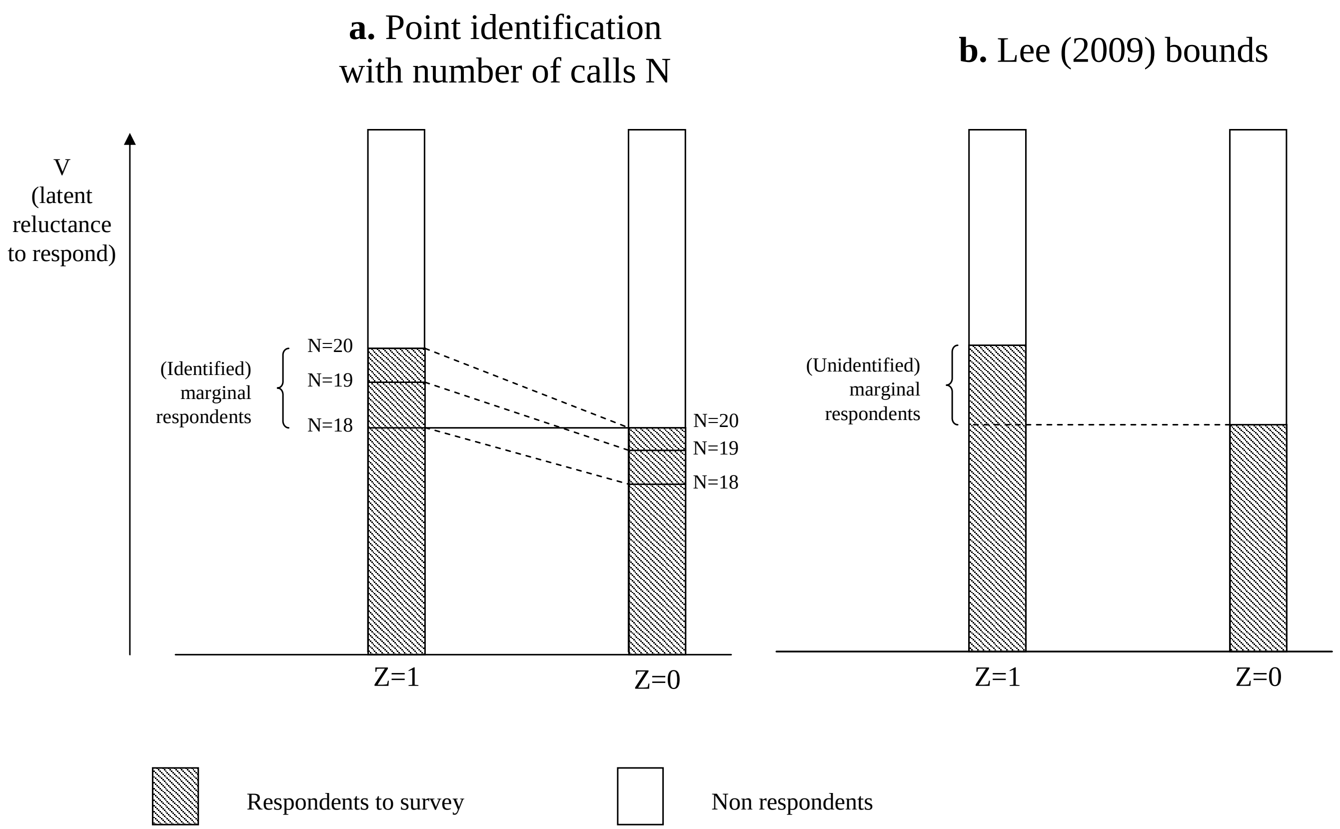


Figure 4.2: Response rates by assignment according to the number of phone calls

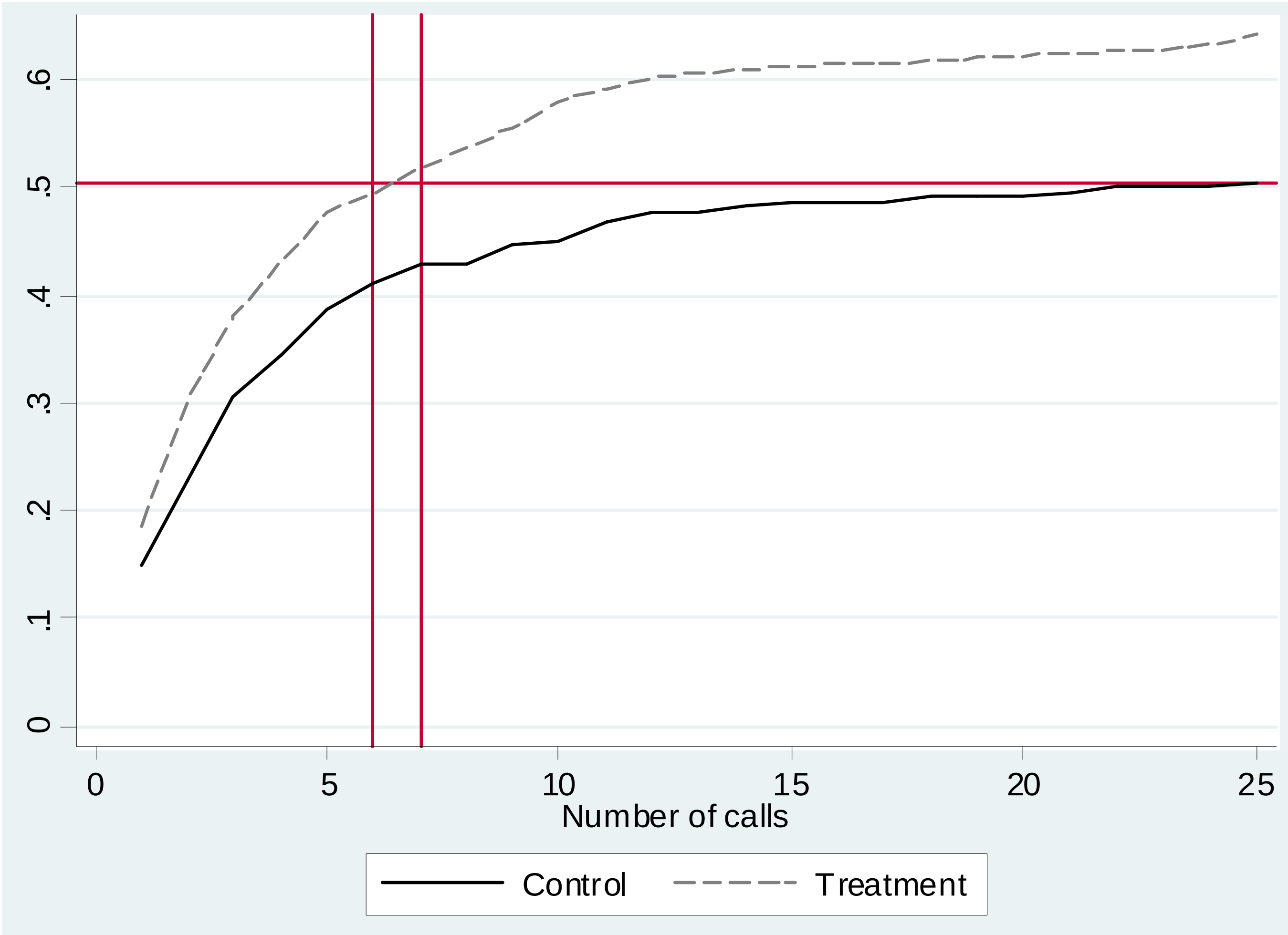


Table 4.1: Population sizes

	Full sample	Respondents to phone survey
Sample size	798	493
Initial population size	4324.82	2478.09
Weighted response rate	1.00	0.57

Note : Response rate is computed using the sampling weights of the telephone survey.

Table 4.2: Program impact on exit from the unemployment registers without correcting for sample selection

	Full sample	Respondents to phone survey
Treatment	0.096	0.136
	(0.040)	(0.054)
N	798	493

Note : Linear probability model. Observations weighted according to the sampling design of the telephone survey. Robust standard errors are below the effects in parenthesis.

Table 4.3: Exit from the unemployment registers depending on the response status in the phone survey

	Non respondents (a)	Respondents (b)	All (c)	Difference (b)-(a)	Difference (c)-(a)	p value (c)-(a)
Control group	0.523	0.378	0.450	-0.145	-0.072	0.031
	(0.048)	(0.046)	(0.034)	(0.067)	(0.033)	.
Treatment group	0.602	0.515	0.546	-0.088	-0.031	0.063
	(0.038)	(0.028)	(0.023)	(0.047)	(0.017)	.

Note : Observations weighted according to the sampling design of the telephone survey. Standard errors are below the effects in parenthesis.

Table 4.4: Impact of assignment on response to phone survey

	Response
Treatment	0.137
	(0.040)
N	798

Note : Linear probability model. Observations weighted according to the sampling design of the telephone survey. Robust standard errors are below the effects in parenthesis.

Table 4.5: Program impact on exit from unemployment records with corrections for sample selection

	Without correction		Horowitz Manski	Lee	Truncation		
Sample	all	respondents			N=6	N=7	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment	0.096	0.136	[-0.356;0.498]	[0.003;0.217]	0.100	0.114	[0.101;.127]
	(0.040)	(0.054)	.	(0.058);(0.053)	(0.084)	(0.11)	.
Size	798	493	493	493	408	423	423

Note : Observations weighted according to the sampling design of the telephone survey. Standard errors are below the effects in parenthesis. Columns (1) and (2) recall table 2. Estimates with standard bounding techniques *à la* Horowitz and Manski and *à la* Lee are in columns (3) and (4). Our truncation procedure are in the last 3 columns. In columns (5) and (6), we abstract from the discreteness of the number of calls and present point estimates under two alternatives for the trimming threshold. In column (7), we take into account the discreteness and present bounds.

4.5 Appendix

4.5.1 Proofs of propositions in the text

4.5.1.1 Proof of proposition 1

Proof. Under assumption 1,

$$\begin{aligned}
E(y|R=1,Z=0,W=w) &= E(y(0)|R=1,Z=0,W=w) \\
&= E(y(0)|\mathbf{1}(V\leq p(w,0)),Z=0,W=w) \\
&= \frac{E(y(0)\mathbf{1}(V\leq p(w,0))|Z=0,W=w)}{\Pr(\mathbf{1}(V\leq p(w,0))=1|Z=0,W=w)} \\
&= \frac{E(y(0)\mathbf{1}(V\leq p(w,0)))}{\Pr(\mathbf{1}(V\leq p(w,0))=1)} \\
&= E(y(0)|V\leq p(w,0)).
\end{aligned}$$

Similarly,

$$E(y|R=1,Z=1,W=w) = E(y(1)|V\leq p(w,1)).$$

This holds for any couples $(w_0,0)$ and $(w_1,1)$. Consequently, if there exists w_0 and w_1 such that $p(w_0,0)=p(w_1,1)=\bar{p}$, then we have:

$$E(y(1)-y(0)|V\leq\bar{p}) = E(y|R=1,W=w_1,Z=1)-E(y|R=1,W=w_0,Z=0),$$

which is equation 4.5 in the text.

□

4.5.1.2 Proof of proposition 3

Proof. By definition, the set $(N(1)\leq w_1)$ is equal to $(V\leq p^*)$. We can decompose $E(y(1)|N(1)\leq w_1)$ depending on wether V is lower than $\bar{p}$ or between $\bar{p}$ and p^* :

$$\begin{aligned}
E(y(1)|N(1)\leq w_1) &= E(y(1)|V\leq\bar{p})\Pr(V\leq\bar{p}|N(1)\leq w_1)... \\
&+ E(y(1)|\bar{p}<V\leq p^*)\Pr(\bar{p}<V\leq p^*|N(1)\leq w_1)
\end{aligned}$$

By manipulating the previous equation we get:

$$E(y(1)|V\leq\bar{p}) = \frac{E(y(1)|N(1)\leq w_1)-E(y(1)|\bar{p}<V\leq p^*)\Pr(\bar{p}<V\leq p^*|N(1)\leq w_1)}{\Pr(V\leq\bar{p}|N(1)\leq w_1)}$$

As $E(y(1)|\bar{p}<V\leq p^*)$ is not observed, we set it to the lower value of y and to its upper bound to obtain our upper and lower bound. □

4.5.2 Estimation and inference of the truncation model

For notation convenience, denote Δ^R the treatment effect identified in proposition 2 ($\Delta^R = E(y(1) - y(0)|V \leq \bar{p})$). In addition to proposition 2 assumption, we assume $\bar{p}$ is actually attained in the control group. Estimation and inference results are inspired by Lee (2009).

First we define the estimates as sample analogs to the parameters defined in proposition 2.

$$\begin{pmatrix} \widehat{\Delta^R} = \frac{\sum YZ\mathbf{1}(N \leq n_{\widehat{\bar{p}}})}{\sum Z\mathbf{1}(N \leq n_{\widehat{\bar{p}}})} - \frac{\sum YR(1-Z)}{\sum R(1-Z)} \\ n_{\widehat{\bar{p}}} = \min n : \frac{\sum Z\mathbf{1}(N \leq n_{\widehat{\bar{p}}})}{\sum Z} \geq \widehat{\bar{p}} \\ \widehat{\bar{p}} = \frac{\sum R(1-Z)}{\sum 1-Z} \end{pmatrix}$$

Second we verify consistency by showing that the estimator solves a well defined GMM problem and applying theorem 2.6 of Newey and McFadden (1986). To do so, we need one additional assumption, i.e. N has bounded support. It is sufficient to prove consistency of $\mu_0 = E(Y|Z = 1, N \leq n_{\bar{p}})$. Denote $\theta_0' = (\mu_0, n_{\bar{p}_0}, \bar{p}_0)'$ the true value of the parameters vector and $d' = (Y, Z, N)'$ the data. Define the moment function $g(d, \theta)$:

$$g(d,\theta) = \begin{pmatrix} (Y-\mu)Z\mathbf{1}(N \leq n_{\bar{p}}) \\ (\mathbf{1}(N \leq n_{\bar{p}}) - \bar{p})Z \\ (\mathbf{1}(N \leq w) - \bar{p})(1-Z) \end{pmatrix}$$

Recall that w is the maximum number of attempts, such that $R = \mathbf{1}(N \leq w)$. The estimator μ_0 is the solution to $min_\theta(\sum g(d, \theta))'(\sum g(d, \theta))$.

Third we verify asymptotic normality by applying theorem 7.2 of Newey and MacFadden (1986). We define $g_0(\theta) = E(g(d, \theta))$ and $\widehat{g_n}(\theta) = n^{-1} \sum g(d, \theta)$. We also define G the derivative of $g_0(\theta)$ at $\theta = \theta_0$. We can verify the assumptions of theorem 7.2 [14] and obtain that the asymptotic variance is $V = G^{-1}\Sigma(G^{-1})'$ where Σ is the asymptotic variance of $\widehat{g_n}(\theta)$. Σ is equal to:

$$\Sigma = \begin{pmatrix} E((Y-\mu_0)^2\mathbf{1}(N \leq n_{\bar{p}_0})|Z=1)E(Z) & 0 & 0 \\ 0 & \bar{p}_0(1-\bar{p}_0)E(Z) & 0 \\ 0 & 0 & \bar{p}_0(1-\bar{p}_0)E(1-Z) \end{pmatrix}$$

Define $f(.)$ as the density of N conditional on $Z = 1$. Then G is equal to:

$$G = \begin{pmatrix} -\bar{p}_0E(Z) & E(Y-\mu_0|Z=1, N=n_{\bar{p}_0})f(n_{\bar{p}_0})E(Z) & 0 \\ 0 & f(n_{\bar{p}_0})E(Z) & -E(Z) \\ 0 & 0 & -E(1-Z) \end{pmatrix}$$

and its inverse G^{-1} is :

[14] the only difficulty is stochastic equicontinuity

$$G^{-1} = \frac{1}{\bar{p_0} f(n_{\bar{p_0}})(E(Z))^2 E(1-Z)} \cdots$$
$$\begin{pmatrix} -f(n_{\bar{p_0}})E(Z)E(1-Z) & Ef(n_{\bar{p_0}})E(Z)E(1-Z) & -Ef(n_{\bar{p_0}})(E(Z))^2 \\ 0 & \bar{p_0}E(Z)E(1-Z) & -\bar{p_0}(E(Z))^2 \\ 0 & 0 & -\bar{p_0}f(n_{\bar{p_0}})(E(Z))^2 \end{pmatrix}$$

where $E = E(Y - \mu_0 | Z = 1, N = n_{\bar{p_0}})$.
Hence the upper left term of the variance matrix is the sum of three terms :

$$\begin{aligned} V(1,1) &= \frac{Var(Y|Z=1, N \leq n_{\bar{p_0}})}{\bar{p_0}E(Z)} + \ldots \\ &\quad \frac{(1-\bar{p_0})(E(Y-\mu_0|Z=1, N=n_{\bar{p_0}}))^2}{\bar{p_0}E(Z)} + \ldots \\ &\quad \frac{(E(Y-\mu_0|Z=1, N=n_{\bar{p_0}}))^2}{E(1-Z)} \end{aligned}$$

The first term (V^Y) is the usual variance of the mean estimator when there is no uncertainty concerning the trimming procedure. The second term (V^N) reflects the fact that once the fraction to be trimmed is known there is still uncertainty about the right number of calls under which the sample should be trimmed. The third term (V^P) is the part of the variance of the estimator due to uncertainty about the true fraction to be trimmed.

To sum up, we have shown that $\sqrt{n}\left(\widehat{\Delta^R} - \Delta_0^R\right) \rightarrow N(0, V^Y + V^N + V^P + V^C)$ in distribution, where V^Y, V^N, V^P are defined just above and V^C is the variance of the conditional mean in the control group: $V^C = \frac{Var(Y|Z=0, N \leq w)}{E(1-Z)p_0}$.

4.5.3 Extension to non compliance

In this appendix, we extend the previous results to the case where compliance is imperfect. We consider the potential outcome framework with random assignment to treatment and imperfect compliance of Angrist, Imbens, and Rubin (1996). $Z \in \{0,1\}$ is the variable related to assignment and $T \in \{0,1\}$ is the final treatment status. The potential treatment variables are $T(0)$ and $T(1)$ (corresponding to $Z = 0$ or $Z = 1$, respectively). Potential outcomes are $y(t,z)$, with $t \in \{0,1\}$ and $z \in \{0,1\}$. We consider the usual set of assumptions of the Angrist, Imbens and Rubin model:

Assumption 2.

1. *SUTVA*

2. *(Monotonicity):*

$$T(1) \geq T(0)$$

3. *(Exclusion)*
$$y(T) = y(T(Z)) \equiv \tilde{y}(Z)$$

4. *(Independence)*

$$Z \perp y(1), y(0), T(1), T(0)$$

(Note that we changed notation for the sake of readability: $y(0)$ and $y(1)$ now denote potential outcome under the different treatment statuses; $\tilde{y}(0)$ and $\tilde{y}(1)$ correspond to potential outcomes under the different assignment statuses that were noted $y(0)$ and $y(1)$ above.) It is well known that under this set of assumptions, the usual Wald estimator identifies the local average treatment effect on compliers (LATE):

$$E(y(1) - y(0)|T(1) - T(0) = 1) = \frac{E(y|Z = 1) - E(y|Z = 0)}{E(T|Z = 1) - E(T|Z = 0)}.$$

We now consider non response. We extend assumption 1 to account for imperfect compliance.

Assumption 3.

1. *(Latent variable threshold-crossing response model):*

$$R = \mathbf{1}(V < \tilde{p}(W, Z)),$$

2. *(Independence):*

$$\begin{aligned} W, Z &\perp \tilde{y}(0), \tilde{y}(1), \tilde{N}(0), \tilde{N}(1), T(0), T(1), V \\ Z &\perp W \end{aligned}$$

Proposition 4. ***Identification with the actual number of calls leading to response under imperfect compliance.*** *Under assumption 2 and 3 and a binary outcome, $E(y(1) - y(0)|V \leq \bar{p}, T(1) - T(0) = 1)$ is identified from the observation of y, T, Z and N:*

$$E(y(1)-y(0)|V \leq \bar{p}, T(1)-T(0) = 1) = \frac{E(y|N \leq w_1, Z = 1) - E(y|N \leq w_0, Z = 0)}{E(T|N \leq w_1, Z = 1) - E(T|N \leq w_0, Z = 0)},$$

with w_0, w_1 such that

$$\Pr(N \leq w_1|Z = 1) = \Pr(N \leq w_0|Z = 0) = \bar{p}.$$

Proof. Under assumption 3, Proposition 2 applies:

$$E(\tilde{y}(1) - \tilde{y}(0)|V \leq \bar{p}) = E(y|N \leq w_1, Z = 1) - E(y|n \leq w_0, Z = 0) \quad (4.12)$$
$$E(T(1) - T(0)|V \leq \bar{p}) = E(T|N \leq w_1, Z = 1) - E(T|n \leq w_0, Z = 0) \quad (4.13)$$

(note that $y = Z\tilde{y}(1) + (1-Z)\tilde{y}(0)$ and $N = Z\tilde{N}(1) + (1-Z)\tilde{N}(0)$.)

By the law of iterated expectations,

$$\begin{aligned} E(\tilde{y}(1)-\tilde{y}(0)|V \leq \bar{p}) \quad = \quad & E(y(1)-y(0)|V \leq \bar{p}, T(1)-T(0)=1) \times \Pr(T(1)-T(0)=1|V \leq \bar{p}) \\ & +0 \times \Pr(T(1)-T(0)=0|V \leq \bar{p}) \\ & +E(y(1)-y(0)|V \leq \bar{p}, T(1)-T(0)=-1) \times \Pr(T(1)-T(0)=-1|V \leq \bar{p}) \end{aligned}$$

In the absence of defiers ($T(1) \geq T(0)$), the last term is 0. Therefore

$$E(y(1)-y(0)|V \leq \bar{p}, T(1)-T(0)=1) \quad = \quad \frac{E(\tilde{y}(1)-\tilde{y}(0)|V \leq \bar{p})}{E(T(1)-T(0)|V \leq \bar{p})} \tag{4.14}$$

Combining equations 4.12, 4.13, and 4.14 yields the result.

□

Proposition 4 implies that $E(y(1)-y(0)|V \leq \bar{p}, T(1)-T(0)=1)$ can be estimated using the standard Wald estimator, after truncating the sample following the order of the number of phone calls needed, up to the point where the same share of the initial population is represented in the treatment and in the control group. Once this truncation is done, control and treatment respondents are statistically identical, and the standard IV argument applies.

4.5.4 Adjustment of Lee (2009) bounds when the outcome is binary

In a context of sample selection, Lee (2009) shows the implications of assuming that selection into the sample is monotonic. In our setting, this amounts to assuming the monotonicity of response behavior in addition to the independence restriction due to random assignment:

Assumption 4.

1. *(Independence)*

$$Z \quad \perp \quad y(1), y(0), R(1), R(0)$$

2. *(Response monotonicity)*

$$R(1) \geq R(0).$$

Intuitively, the Lee (2009) bounds build on the idea that treatment assignment only impacts response behavior by bringing in new respondents: all those who respond when assigned to treatment also respond when assigned to control. It is therefore possible to put bounds on the treatment effect by trimming the upper and lower tail of the outcome distribution according to the number of excess respondents. With a continuous y, this implies estimating the corresponding quantiles of the outcome distribution. In our case, however, quantiles are not useful as y is either 0 or 1. We therefore propose what can be viewed as "hybrid" of the Horowitz and Manski (2000) and Lee (2009) in that it assumes both a bounded support for the outcome and makes a monotonicity restriction:

Proposition 5. ***Identification under monotonicity with a binary outcome.*** *If assumption 4 holds and the outcome is binary ($y(0), y(1) \in \{0,1\}$), then $E(y(1) - y(0)|R(1) = 1)$ belongs to an identifiable interval whose lower and upper bounds are:*

$$\begin{aligned}\Delta^{LB} &= \frac{E(yR|Z=1) - E(yR|Z=0)}{E(R|Z=1)} - \frac{E(R|Z=1) - E(R|Z=0)}{E(R|Z=1)} \\ \Delta^{UB} &= \frac{E(yR|Z=1) - E(yR|Z=0)}{E(R|Z=1)}.\end{aligned}$$

Proposition 5 makes it clear that compared to the Manski and Horowitz (2000) bounds, the length of the identifiable interval, $\frac{E(R|Z=1)-E(R|Z=0)}{E(R|Z=1)}$, depends primarily on the difference in response rates, $E(R|Z=1) - E(R|Z=0)$, and not on their sum.

Proof.

$$\begin{aligned}E(y(1) - y(0)|R(1) = 1) &= \frac{E(y(1)R(1) - y(0)R(1))}{E(R(1))} \\ &= \frac{E(y(1)R(1) - E(y(0)R(0)) - E(y(0)(R(1) - R(0)))}{E(R(1))} \\ &= \frac{E(yR|Z=1) - E(yR|Z=0)}{E(R|Z=1)} - \frac{E(y(0)(R(1) - R(0)))}{E(R|Z=1)}.\end{aligned}$$

The only term that is not identified from the data is $E(y(0)(R(1) - R(0)))$. It can however be bounded:

$$\begin{aligned}E(y(0)(R(1) - R(0))) &= E(y(0)|R(1) - R(0) = 1)\Pr(R(1) - R(0) = 1) \\ &- E(y(0)|R(1) - R(0) = -1)\Pr(R(1) - R(0) = -1) \\ &= E(y(0)|R(1) - R(0) = 1)(\Pr(R=1|Z=1) - \Pr(R=1|Z=0)),\end{aligned}$$

due to the monotonicity condition. Taking into account that $y(0) \in \{0,1\}$, this yields Proposition 5.

□

Chapter 5

Labor Market Policy Evaluation in Equilibrium: Some Lessons of the Job Search and Matching Model[1]

Contents

5.1 Introduction

Most policy evaluations are based on comparing the behavior of participants and non participants in the policy. But the differences in outcome between the treatment group and the control group do estimate the policy mean impact only if the outcomes of the control group are not influenced by the policy, the so-called 'no-interference' (Rubin (1978)) or 'stable unit treatment value' (Angrist, Imbens, and

[1]This chapter is largely based on common work with Pierre Cahuc.

Rubin (1996)) assumption. However, the policy may have equilibrium effects that extend to the untreated as well. For instance, Heckman, Lochner, and Taber (1998) strikingly illustrate this point in the context of education policies. This issue, which is discussed in a broader perspective in the survey of Meghir (2006), is particularly relevant to the evaluation of labor supply based policies (such as increasing incentives or monitoring the unemployed). First, they generally aim at increasing the overall number of filled jobs, which depends on the interactions between aggregate labor supply and labor demand. Second, these policies may induce displacement effects: treated persons may crowd out the untreated because they compete for the same jobs.

Although they have long been recognized, these questions have received limited attention to date. Davidson and Woodbury (1993) and Calmfors (1994) are early contributions. More recently, Lise, Seitz, and Smith (2005) study the equilibrium effects of the Self-Sufficient Project incentive program in Canada. They calibrate an equilibrium model of the labor market so that, when used in partial equilibrium, the model matches the effect of the program estimated by direct comparison of treated and untreated. When equilibrium effects are simulated, the impact of the Self-Sufficient Project is far lower. In contrast, Albrecht, van den Berg, and Vroman (2009) find, using a calibrated model, equilibrium effects of a Swedish training program to be stronger than implied by direct comparison. Using a job search and matching model with skilled and unskilled workers, Van der Linden (2005) shows that micro and equilibrium evaluations are likely to differ widely when job search effort and wages are endogenous. When wages are bargained over, raising the effectiveness of or the access to counseling programs pushes wages upwards and leads to lower search effort among nonparticipants. Induced effects can outweigh positive micro effects on low-skilled employment when the response of wages is taken into account.

The equilibrium effects have also been analyzed in empirical evaluations that do not rely on structural models. For instance, the contribution of Blundell, Costa Dias, Meghir, and J. (2004) evaluates the New Deal for Young People in the U.K. This program was piloted in certain areas before it was rolled out nation wide. Moreover, the program has age specific eligibility rules. Blundell, Costa Dias, Meghir and Van Reenen use these area and age based eligibility criteria that vary across individuals of identical unemployment durations to identify the program effects. They find that either equilibrium wage and displacement effects are not very strong or they broadly cancel each other out.

The aim of our paper is to analyze the impact of counseling in the standard matching model of the labor market (Pissarides, 2000). In our specification, counseled unemployed have a constant comparative advantage in the job search.[2] Using

[2] We simply assume that counseling increases the exit rate out of unemployment. Monitoring and sanctions are not explicitly considered here (for an overview, see Boone, Fredriksson, Holmlund, and van Ours (2007)). Counseling programs are very different from long-duration training schemes intended to enhance skills (see Albrecht, van den Berg, and Vroman (2009), Boone, Fredriksson,

this simple model allows us to analyze the consequences of counseling in a dynamic set-up, whereas previous studies are limited to the comparison of steady states. More precisely, we shed some light on three important issues:

(i) What is the true impact of the policy when equilibrium effects are taken into account? The model shows that the true impact of counseling can be very different from what can be concluded when equilibrium effects are neglected even when the treatment group is small. For instance, we find that counseling can increase unemployment when a small proportion of job seekers benefit from counseling, although counseling improves the efficiency of job search. Equilibrium effects rely on the adjustment of wages. The impact of policies on wages has been analyzed in some papers devoted to equilibrium effects of several labor market policies and education policies, in particular since the seminal contribution of Heckman, Lochner and Taber (1998).[3] Our model allows us to analyze precisely the reaction of wages to counseling, as in the paper of Van der Linden (2005).[4]

(ii) What is the impact of the generalization of the policy to a large treatment group? The model shows that there is no simple answer. In particular, the relation between the impact of the policy on unemployment and the size of the treatment group is not necessarily monotonic. Strikingly, in our framework, unemployment increases with the size of the treatment group when a small share of job seekers are treated but diminishes with the size of the treatment group when a sufficiently large share of job seekers are counseled.

(iii) What is the dynamic impact of counseling? Many experiments made to evaluate labor market policies are transitory. Typically, a group of job seekers is selected to benefit from counseling (the treatment group) and the control group will never benefit from counseling. The comparison between the outcomes yields the evaluation of the impact of counseling. Our model allows us to stress that the consequences of permanent and transitory policies can be very different. The difference comes from the reaction of non-counseled job seekers. When the policy is transitory, non-counseled workers do not expect to benefit from counseling in the future. However, when the policy is permanent, the expectation to benefit from counseling in the future induces the non-counseled workers to raise their reservation wage. In our framework, this phenomenon implies that permanent counseling increases unemployment when a small share of job seekers are counseled whereas counseling always decreases unemployment when it is transitory. Accordingly, it can be misleading to conclude that a truly successful transitory policy will remain successful when it becomes permanent.

The paper is organized as follows. The model is presented in section 2. Section 3 is devoted to the impact of counseling in steady state. Transitory dynamics are analyzed in section 4. Section 5 provides concluding comments.

Holmlund, and van Ours (2007), Masters (2000)).

[3] See the survey of Meghir (2006).

[4] Van der Linden assumes that wages are collectively bargained over, whereas we assume an individual bargaining framework, where counseled and non-counseled workers can get different wages.

5.2 The model

We consider a standard matching model *à la* Pissarides (2000) with a continuum of infinitely-lived risk neutral workers. The measure of the continuum is normalized to one. There are two goods: a good produced and consumed, which is the numeraire, and labor. There is a common discount rate r, strictly positive. Time is continuous. Workers can be in three different states: (1) employed, (2) unemployed and counseled, (3) unemployed and not counseled. Upon entering unemployment, workers are not counseled. They then enter into counseled status at a rate $\mu > 0$ and they keep on receiving counseling until they find a job.

There is an endogenous number of jobs. Each job can be either vacant or filled. Filled jobs produce $y > 0$ units of the numeraire good per unit of time, whereas vacant jobs cost c per unit of time. Filled jobs are destroyed with probability $\lambda > 0$ per unit of time.

Vacant jobs and unemployed workers (the only job seekers, by assumption) are brought together in pairs through an imperfect matching process. This process is represented by the customary matching function, which relates total contacts per unit of time to the seekers on each side of the market. Let us denote by u_n and u_c the number of non-counseled and counseled unemployed workers respectively. In our set-up, the only potential effect of counseling is to increase the arrival rate of job offers to the counseled unemployed workers. Let us normalize to one the number of efficiency units of job search per unit of time of each non-counseled unemployed worker. Counseled unemployed workers are assumed to produce a different number of efficiency units of search, denoted by $\delta \geq 1$.[5] In this setting, the number of efficiency units of job search per unit of time amounts to $s = u_n + \delta u_c$. It should be noted that empirical studies do not systematically find a positive impact of counseling on the entry rate into employment. For instance, Van den Berg and van der Klaauw (2006) find that counseling and monitoring do not affect the exit rate to work in the Dutch unemployment insurance system at the end of the 1990s. Crépon, Dejemeppe, and Gurgand (2005) find a significant positive impact of counseling in France over the period 2002-2004. Here, we simply assume that counseling has a positive impact on the entry rate into work at the individual level in order to analyze the equilibrium effects of counseling.

The number of employer-worker contacts per unit of time is given by $M(s, v) \geq 0$, where $v \geq 0$ denotes the number of job vacancies and M is the matching function, twice continuously differentiable, increasing, concave in both of its arguments, and linearly homogeneous. Linear homogeneity of the matching function allows us to express the probability per unit of time for a vacant job to meet an unemployed worker as a function of the labor market tightness ratio, $\theta = v/s$. A vacant job meets on average $M(s, v)/v = q(\theta)$ unemployed workers per unit of time, with $q'(\cdot) < 0$. Similarly, the rate at which counseled and non counseled unemployed job seekers can meet jobs is $\delta\theta q(\theta)$ and $\theta q(\theta)$ respectively.

[5] Pissarides (1979) and more recently Cahuc and Fontaine (2009) provide models that explicitly represent how the employment agency can increase the efficiency of matching.

Parameter δ is estimated by econometricians who evaluate the impact of counseling by comparing the exit rate out of unemployment of counseled workers and the exit rate out of unemployment of non-counseled workers assuming that the arrival rate of job offers to the non-counseled workers is not influenced by counseling. Henceforth, we assume that δ has been correctly evaluated in this way. The model allows us to analyze the impact of counseling on the non-counseled workers and on labor market equilibrium.

5.2.1 Job creation

Let J_c and J_n be the present-discounted value of expected profit from an occupied job with a counseled worker and a non-counseled worker respectively. Let V denote the present-discounted value of expected profit from a vacant job. V satisfies

$$rV = -c + q(\theta)\left[\alpha J_c + (1-\alpha)J_n - V\right] + \dot{V},$$

where $\dot{V}$ denotes the time derivative of V and

$$\alpha = \frac{\delta u_c}{\delta u_c + u_n} \tag{5.1}$$

stands for the probability to meet a counseled worker. The free entry condition for the supply of vacant jobs is $V = 0$ at any date, implying that

$$\frac{c}{q(\theta)} = \alpha J_c + (1-\alpha)J_n. \tag{5.2}$$

Let us denote by w_c and w_n the wage of a counseled worker and of a non-counseled worker respectively. The asset value of a job filled with a counseled worker, J_c, satisfies

$$rJ_c = y - w_c + \lambda(V - J_c) + \dot{J}_c. \tag{5.3}$$

Similarly, the asset value of a job filled with a non-counseled worker, J_n, satisfies

$$rJ_n = y - w_n + \lambda(V - J_n) + \dot{J}_n. \tag{5.4}$$

At this stage, it can be shown that the impact of counseling on the arrival rate of job offers to the non-counseled depends on the wages w_c and w_n.

5.2.2 The impact of counseling when wages are exogenous

Let us assume for a while that wages w_c and w_n are exogenous. Then, equations (5.3) and (5.4), which define the asset value of filled jobs, imply that $J_c = (y - w_c)/(r+\lambda)$ and $J_n = (y - w_n)/(r+\lambda)$.[6] Substituting these expressions into the free entry condition (5.2) yields

$$\frac{c(r+\lambda)}{q(\theta)} = y - \left[\alpha w_c + (1-\alpha)w_n\right]. \tag{5.5}$$

[6] Since wages are constant in both equations the solution must satisfy $\dot{J}_c = \dot{J}_n = 0$ to be compatible with finite values of J_n and J_c.

From equation (5.1), it turns out that increases in the share of counseled workers increase the probability α that firms meet counseled workers.[7] Then, equation (5.5) shows that increases in α reduce labor market tightness (and then the exit rate out of unemployment of the non-counseled, equal to $\theta q(\theta)$) if the wage of counseled workers is higher than the wage of the non-counseled. In this case, increases in the share of counseled workers raise the proportion of high paid workers. Then, expected profits decrease and firms post fewer job vacancies. If counseled workers get lower wages than non-counseled workers, we get the opposite result: counseling increases labor market tightness. When wages are identical, labor market tightness is independent of the share of counseled workers. This may be the case when there is a minimum wage that is binding for both counseled and non-counseled workers.

The analysis of the case where wages are exogenous allows us to stress the role played by wage adjustment. In our simple search and matching model where workers are ex-ante identical, counseling may have an impact on labor market tightness, and then on the arrival rate of job offers to the non-counseled workers, if it induces wage differentials between the counseled and the non-counseled.

5.2.3 Wage bargaining

Let us now suppose that wages are bargained over. Wage negotiation sets wages that can be renegotiated by mutual agreement only. This means that neither party can oblige the other to renegotiate except if she has a credible threat to do so. In other words, a party can force the other to renegotiate if her outside option yields higher gains than job continuation at the current wage. In our setup, the employer can trigger a renegotiation only if the expected profits of the filled job, at the current wage, are smaller than the expected profits that she would get by firing the worker. In the same manner, the employee can trigger a renegotiation only if she prefers to quit her job rather than go on working at the current wage. As stressed by Malcomson (2011) and Cahuc, Postel-Vinay, and Robin (2006), this assumption is in line with the legal rules in most OECD countries, which state that an offer to modify the terms of a contract does not constitute a repudiation. Accordingly, a rejection of the offer to renegotiate by either party leaves the preexisting terms in place, which means that the job continues under those terms if the renegotiation is refused. In our framework, where the productivity y is constant over time and where the equilibrium value of job vacancies is equal to zero, employers cannot trigger renegotiations because they always make positive profits with filled jobs at the current wage. The employees are also unable to renegotiate the wage. When they are matched with a new employer, they are counseled and they continue to be counseled if they do not reach an agreement that allows them to be employed.

[7] From equation (5.1), we have

$$\alpha = \frac{\delta}{\delta + (u_n/u_c)}$$

so that increases in the share of counseled workers, which reduce, by definition, the ratio u_n/u_c, necessarily increase α.

Once they have accepted their job, they are not counseled further if they enter into unemployment. Therefore, for the employees, the outside option on continuing jobs is smaller than that on new jobs. This implies that they cannot be in position to renegotiate their wage. Finally, in our framework, the assumption of renegotiation by mutual agreement implies that the wage remains constant over the full duration of the job.[8]

Let us define the workers' returns when employed and unemployed in order to derive the outcome of the wage bargaining. The present-discounted value of the expected income stream of, respectively, a counseled and a non-counseled unemployed worker, is denoted by U_c and U_n. The present-discounted value of the expected income stream of employees who found a job while counseled is denoted by W_c. The present discounted value of the employees who obtained a job without being counseled is denoted by W_n. All unemployed workers enjoy some instantaneous return z which includes unemployment benefits and the imputed return of leisure. The non-counseled workers exit unemployment at rate $\theta q(\theta)$ and enter into counseling at rate μ. The counseled ones exit unemployment at rate $\delta\theta q(\theta)$. Hence U_n, U_c, W_n and W_c satisfy

$$\begin{aligned} rU_n &= z + \mu(U_c - U_n) + \theta q(\theta)(W_n - U_n) + \dot{U}_n, && (5.6)\\ rU_c &= z + \delta\theta q(\theta)(W_c - U_c) + \dot{U}_c, && (5.7)\\ rW_n &= w_n + \lambda(U_n - W_n) + \dot{W}_n, && (5.8)\\ rW_c &= w_c + \lambda(U_n - W_c) + \dot{W}_c. && (5.9) \end{aligned}$$

We assume that the wage bargaining outcome yields a share β of the surplus of the job to the worker. The surplus of a job filled by a previously counseled worker is

$$S_c = W_c - U_c + J_c - V.$$

The surplus of a job filled by a worker who did not benefit from counseling is

$$S_n = W_n - U_n + J_n - V.$$

The surplus sharing rule reads

$$W_i - U_i = \beta S_i, J_i - V = (1-\beta)S_i, i = c, n. \tag{5.10}$$

The outcome of the wage bargaining being defined, it becomes possible to derive the set of equations that defines the value of endogenous variables in equilibrium.

[8]This is always true in steady state. However, if there is a large positive non-anticipated shock on μ, it is possible that the outside option of employees becomes larger than the current value of their job. In that case, they renegotiate their wage. But this renegotiation has no effect either on labor tightness or on job destruction, which is exogenous in our model. Accordingly, the dynamics of unemployment remain the same whether wages are renegotiated or not.

5.2.4 Labor market equilibrium

Using the sharing rule, the definitions of the surpluses and equations (5.6) through (5.9) we can write

$$(r+\lambda)S_c - \dot{S}_c = y - z - \theta q(\theta)\delta\beta S_c - \lambda\Delta, \tag{5.11}$$
$$(r+\lambda)S_n - \dot{S}_n = y - z - \theta q(\theta)\beta S_n - \mu\Delta, \tag{5.12}$$

where $\Delta = W_c - W_n > 0$ satisfies

$$(r+\mu)\Delta - \dot{\Delta} = \theta q(\theta)\beta(\delta S_c - S_n). \tag{5.13}$$

Equations (5.11), (5.12) and (5.13) comprise four unknown variables: S_c, S_n, θ and Δ. Using the free entry condition (5.2) together with the sharing rule (5.10), we obtain a relation between labor market tightness θ and the surpluses which involves two more unknowns u_n and u_c :

$$\frac{c}{q(\theta)} = (1-\beta)\left(\frac{u_n}{\delta u_c + u_n}S_n + \frac{\delta u_c}{\delta u_c + u_n}S_c\right). \tag{5.14}$$

Then, the relations between labor market tightness and the unemployment rates are derived from the law of motion of u_n and u_c, which read

$$\dot{u}_n = \lambda(1-u_n-u_c) - \mu u_n - \theta q(\theta)u_n \tag{5.15}$$
$$\dot{u}_c = \mu u_n - \delta\theta q(\theta)u_c. \tag{5.16}$$

Finally, the system of six equations from (5.11) to (5.16) comprises six unknown variables $S_n, S_c, \theta, \Delta, u_n, u_c$.

5.2.5 The impact of counseling on labor market equilibrium with endogenous wages

The analysis of the steady state solution of the system of equations (5.11) to (5.16) allows us to shed light on the impact of counseling on labor market equilibrium. This can be done by looking at the free entry condition (5.14). The left hand side of this equation is the expected cost of a vacant job, which is equal to the instantaneous cost, c, times the average duration $1/q(\theta)$. The expected cost of a vacant job is increasing with labor market tightness θ, because the average duration of vacancies is higher when labor tightness increases. The right hand side of equation (5.14) is the expected profit of a match between a vacant job and a worker. It turns out that the expected profit is equal to the employer's share $(1-\beta)$ of the surplus, times the average value of the surplus. The average value of the surplus is a convex combination of the surplus of jobs filled with counseled workers, S_c, and of the surplus of jobs filled with non-counseled workers. S_n. Equations (5.11) and (5.12) show that the surplus of jobs filled with counseled workers is smaller than the surplus of jobs filled with non-counseled workers.[9] The surplus of jobs filled with counseled

[9] When $\delta > 1$, S_n is necessarily larger than S_c. Suppose that this is not the case, so that $S_c \geq S_n$,then $\delta S_c > S_n$. From equations (5.11), (5.12) and (5.13), we obtain the following expression:

workers is smaller because the reservation wage of counseled workers is higher than the reservation wage of non-counseled workers.

With this property in mind, it can easily be understood how counseling can reduce labor market tightness by looking at the free entry condition (5.14). First, an increase in the proportion of counseled workers raises the probability that vacant jobs are matched with counseled workers who yield filled jobs with relative low surplus. This reduces the expected profits of filled jobs and then induces firms to create fewer job vacancies. Second, everything else being equal, an increase in the proportion of counseled workers decreases the value of the surplus of jobs filled with non-counseled workers because it improves their outside option. This effect also contributes to reduce expected profits and then labor market tightness. Third, everything else being equal, the value of the surplus of jobs filled with counseled workers increases when there is more counseling. If the two first effects dominate, which is the case for simulations done with a large range of relevant values of the parameters, counseling induces fewer job offers to the non-counseled.[10] In appendix B, we show that the two first effects dominate when the share of counseled persons is small if the matching function takes the form $q(\theta) = \theta^{1/2}$ that will be used in our calibration exercises.

Once the effect of counseling on labor market tightness is known, it is possible to look at its impact on the steady state unemployment rate, $u = u_n + u_c$, which can be computed from equations (5.15) and (5.16). Let us denote the steady state value of the unemployment rate as a function of labor market tightness θ and the entry rate into counseling μ as

$$u(\theta, \mu) = \frac{\lambda\left[\delta\theta q(\theta) + \mu\right]}{\lambda\left[\delta\theta q(\theta) + \mu\right] + \delta\theta q(\theta)\left[\mu + \theta q(\theta)\right]}.$$

This expression of the unemployment rate allows us to write its derivative with respect to the entry rate into counseling:

$$\frac{\mathrm{d}u(\theta, \mu)}{\mathrm{d}\mu} = \frac{\partial u(\theta, \mu)}{\partial \mu} + \frac{\partial u(\theta, \mu)}{\partial \theta}\frac{\mathrm{d}\theta}{\mathrm{d}\mu} \tag{5.17}$$

It can easily be checked that $\partial u(\theta, \mu)/\partial\mu < 0$, and that $\partial u(\theta, \mu)/\partial\theta < 0$. The interpretation of the sign of these partial derivatives is straighforward. First, an increase

$$S_c - S_n = -\frac{\theta q(\theta)\beta}{r + \mu}(\delta S_c - S_n)$$

which implies that $S_c < S_n$, which is incompatible with the assumption that $S_c \geq S_n$. This enables us to conclude that $S_c < S_n$.

[10]It is worth noting that the analysis of the properties of the equilibrium with endogenous wages is in line with the results obtained above in the equilibrium with exogenous wages. It has been shown that increasing the share of counseled workers decreases labor market tightness if the wage of counseled workers is higher than the wage of non-counseled workers when wages are exogenous. The negotiated wage of counseled workers is indeed higher than that of the non-counseled. This is the consequence of $S_n > S_c$. Using the bargaining solution $J_i = (1-\beta)S_i$, $i = c, n$ and equations profits, $S_c < S_n$ implies $J_c < J_n$ which implies that $w_c = y-(r+\lambda)(1-\beta)S_c > w_n = y-(r+\lambda)(1-\beta)S_n$.

in the entry rate into counseling raises the share of unemployed who exit unemployment at a higher rate. The effect on unemployment, everything else being equal, is negative: $\partial u(\theta,\mu)/\partial\mu < 0$. Second, when labor market tightness is increased, the exit rate out of unemployment is higher and unemployment drops: $\partial u(\theta,\mu)/\partial\theta < 0$.

When counseling reduces labor market tightness, the term $\frac{\partial u(\theta,\mu)}{\partial\theta}\frac{\mathrm{d}\theta}{\mathrm{d}\mu}$ in the right hand side of equation (5.17) is positive and the total impact of counseling on steady state unemployment is ambiguous.

5.3 Policy evaluation in steady state

In this section we calibrate the model and we analyze the equilibrium effect of counseling in steady state.

5.3.1 Calibration

The frequency of the model is monthly. The 3 month interest rate is set to 1.2 percent, which makes the monthly discount factor equal to 0.996. We need to specify the matching function: $q(\theta) = q_0\theta^{\sigma}$ We choose a conservative value for the elasticity $\sigma = 0.5$. The bargaining power β is set equal to σ to ensure that the Hosios condition is fulfilled (in the model without counseling). We aim to reproduce features of the French labour market (means are taken from 2000 to 2007, which correponds to the last business cycle). The instantaneous return of unemployment, z, is equal to 60 percent of the productivity y, which value is normalized to one. This implies a replacement ratio (z over w) slightly above 60 percent since wages take values around 0.96 in equilibrium. The mean unemployment duration, measured in the Labor Force Survey ("Enquête emploi") between 2004 and 2005,[11] is 1.07 year. The monthly exit rate out of unemployment consistent with this mean unemployment duration is $e_{ss} = 7.80$ percent. The overall unemployment rate averaged 9.5 percent over the same time period.

To compute the baseline equilibrium, we assume that there is no counseling so that $\mu = u_c = 0$ and $u_n = 0.095$. The separation rate is thus $\lambda = 0.8$ percent. The cost of posting a vacancy is set to be roughly one third of a period of production $c = 0.3y$.

The value of parameter q_0 of the matching function is determined by the following relation[12]

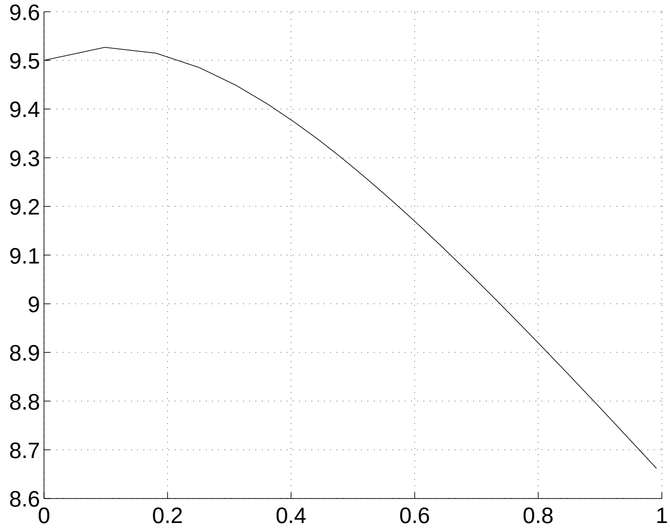
$$q_0 = \frac{c}{y-z}\left(e_{ss}\right)^{\sigma/(1-\sigma)}\left[\left(r+\lambda+\beta e_{ss}\right)\frac{1}{1-\beta}\right]^{1-\sigma}. \tag{5.18}$$

[11] The measure of unemployment duration is the average length of unemployment spells in progress.

[12] This equation is obtained from the free entry condition when there is no counseling, i.e. when $\mu = 0$. In that case, the free entry condition simply reads

$$\frac{c}{q(\theta)} = (1-\beta)\frac{y-z}{r+\lambda+\beta\theta q(\theta)}.$$

Figure 5.1: The relation between the unemployment rate (y-axis) and the share of counseled workers (x-axis).



5.3.2 Policy experiment

In this subsection, we look at the consequences of the introduction of a counseling policy that improves the efficiency of the search activity of counseled workers. We assume that non-counseled workers produce one unit of search per unit of time, so that their arrival rate of job offers amounts to $\theta q(\theta)$. In line with the estimations of Crepon et al. (2005), we assume that the counseled produce $\delta = 1.2$ unit of search per unit of time, so that their arrival rate of job offers is $1.2 \times \theta q(\theta)$.

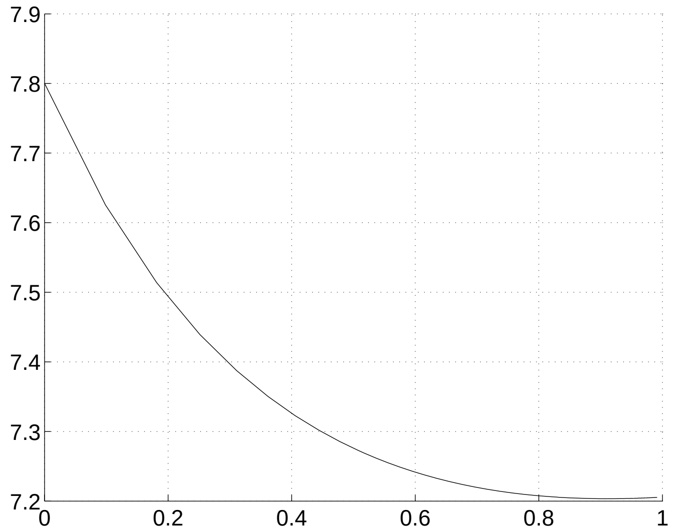
5.3.2.1 The impact of counseling on unemployment

Figure 5.1 displays the relation between the unemployment rate and the share of counseled workers in steady state. It is striking that steady state unemployment increases with the share of counseled workers when this share is small, below 10 percent. This result shows that a naive evaluation, relying on a simple comparison of the outcomes of participants and non-participants that neglects equilibrium effects, can lead to the wrong conclusion that counseling decreases unemployment, especially when the share of counseled workers is small.

Obviously, the negative impact of counseling on unemployment comes from its effect on the arrival of job offers to the non-counseled. Figure 5.2 shows that the arrival rate of job offers to the non-counseled decreases with the share of counseled workers. The drop in the baseline arrival rate of job offers, $\theta q(\theta)$, is the result of two effects. First, there is a decrease in profitability due to the new composition of the unemployed population. Because the counseled get higher wages than the non-counseled,[13] a spread of counseling reduces profitability, and this composition

[13] Recall that the counseled get higher wages because counseling enhances their exit rate out of

Figure 5.2: The relation between the arrival rate of job offers to non-counseled workers (y-axis) and the share of counseled workers (x-axis).



effect hinders job creation. Formally, if we differentiate the free entry condition (5.5), we get:

$$-\frac{c(r+\lambda)q'(\theta)}{(q(\theta))^2}\frac{\partial\theta}{\partial\mu} = -\frac{\partial\alpha}{\partial\mu}(w_c - w_n) - \left[\alpha\frac{\partial w_c}{\partial\mu} + (1-\alpha)\frac{\partial w_n}{\partial\mu}\right].$$

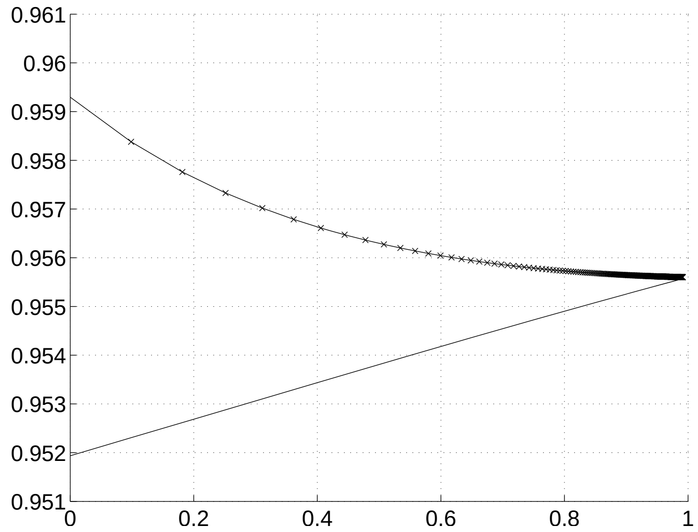
The first term of the right-hand side corresponds to the composition effect. The second effect, which shows up in the second term, comes from the adjustment of wages (see Figure 5.3). The wage of non-counseled workers is pushed upward by counseling because non-counseled workers anticipate that they may benefit from counseling in the future. In contrast, the wage of counseled workers diminishes with the entry rate into counseling. To understand this property, one has to be aware that counseling creates an opportunity cost of accepting job offers: counseled job seekers who find jobs can loose them and will then have to wait a while before benefiting from counseling again. This opportunity cost is higher when the probability of being counseled again, after the accepted job is lost, is lower. Therefore, the opportunity cost to accept a job, and then the negotiated wage, is higher when the entry rate into counseling is smaller.

Finally, the composition effect and the wage effect result in a negative impact of counseling on the baseline arrival rate of job offers, $\theta q(\theta)$, as shown by Figure 5.2.

The decline in the baseline arrival rate of job offers induced by counseling tends to drive the unemployment rate upwards. This effect competes with the direct effect of counseling which makes counseled job seekers leave unemployment faster. When the share of counseled workers is small, the first effect dominates: the share of non-counseled workers who are adversely affected is large and counseled workers

unemployment and then their reservation wage.

Figure 5.3: The relation between the wage of counseled workers (broken line) and the wage of non-counseled workers (continuous line) (y-axis) and the share of counseled workers (x-axis).



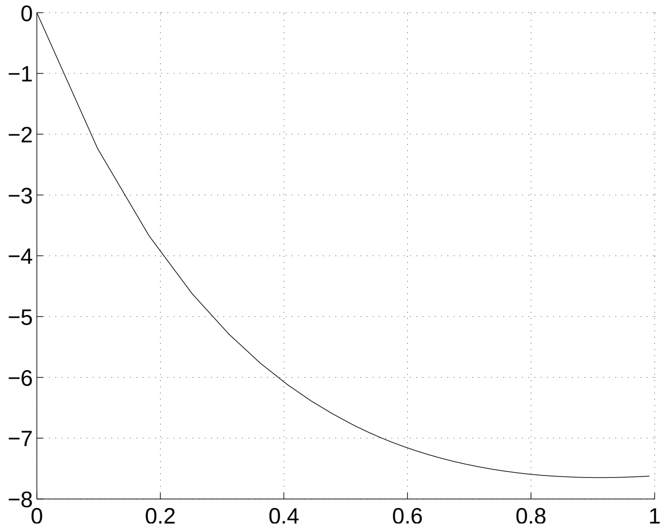
get very high wages. When the share of counseled workers is large, the second effect dominates: even if counseled workers are numerous, they get lower wages than when they are fewer.

5.3.2.2 Evaluation errors

Our model allows us to shed light on the size of the evaluation errors caused by ignoring equilibrium effects. Standard evaluations, relying on a simple comparison of the outcome of the treated and the non treated, can lead to wrong results if the policy induces equilibrium effects which change the baseline arrival rate of job offers $\theta q(\theta)$. The error comes from the choice of wrong counterfactuals when evaluating the impact of the policy: standard evaluations assume that the counterfactual arrival rates of job offers to the non-treated in the absence of the policy are the same as those observed by the econometrician in the presence of the policy.

In our model, the exit rate out of unemployment of counseled job seekers amounts to $\delta\theta q(\theta)$. Non-treated individuals exit unemployment at rate, $\theta q(\theta)$. The effect of the treatment on the treated is usually defined as the ratio between these two exit rates, that is δ. However, this approach yields a naive evaluation of the effects of the treatment to the extent that it does not account for equilibrium effects which may change the value of the arrival rate of job offers to the non-counseled job seekers. To account for such effects one needs to know the exit rate out of unemployment in the absence of counseling, which we denote by $\theta_0 q(\theta_0)$. Then, the effect of the treatment on the treated accounting for equilibrium effects is defined as $\delta\theta q(\theta)/\theta_0 q(\theta_0)$. The error induced by the ignorance of equilibrium effects, expressed as a percentage of the naive evaluation δ, is thus $\left[\theta q(\theta) - \theta_0 q(\theta_0)\right]/\theta_0 q(\theta_0)$. Figure

Figure 5.4: The relation between the error (in percentage of the naive evaluation δ) in the evaluation of the effect of counseling on the exit rate out of unemployment of counseled workers (y-axis) and the share of counseled workers (x-axis).



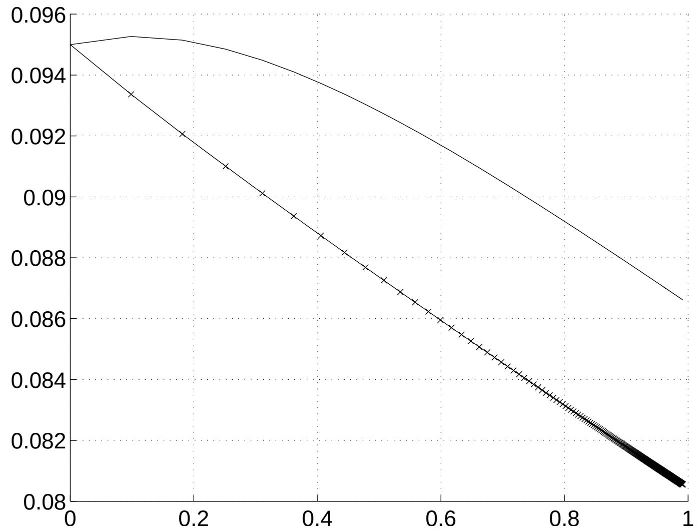
5.4 shows that the naive evaluation leads to an over estimation of the 'true' effect. The absolute error increases with the share of counseled workers. It is equal to 4 percent when the share of counseled workers amounts to 20 percent and reaches nearly 8 percent when the share goes to one.

Another error can occur when simulating the consequence of the spread of the policy to all workers. Looking at this error is important to the extent that some policy makers think that policies should first be evaluated at a small scale before being generalized if their evaluations are favorable. This idea is right only if equilibrium effects are properly taken into account. Ignoring such effects can lead to false conclusions, because it is wrong to simulate the impact of the generalization of counseling to all job seekers with the assumption that the arrival of job offers remains unchanged. We can shed light on this type of error by looking at the difference between the true value of the unemployment rate, denoted by u^*, and the value of the unemployment rate, denoted by $\tilde{u}$, computed when it is assumed that the baseline arrival rate remains unchanged, equal to $\theta_0 q(\theta_0)$. Figure 5.5 plots the true unemployment rate, u^*, (continuous line) and the unemployment rate computed without accounting for equilibrium effects, $\tilde{u}$.

5.4 Policy evaluation and dynamic adjustment

Up to now, we have analyzed the impact of counseling on labor market equilibrium in steady state. It is also important to keep in mind that most labor market policies induce dynamic adjustments that take time. Our model allows us to study the dynamic path of the endogenous variables. We consider three policy experiments

Figure 5.5: The relation between the unemployment rate (true unemployment rate: continuous line, equilibrium unemployment rate computed without accounting for equilibrium effects: broken line) and the share of counseled workers (x-axis).



that differ in the proportion of people being counseled. In the baseline scenario the entry rate into counseling, μ, is equal to 5 percent. There is also a 'light' scenario, where μ is equal to 1 percent, and an 'intensive' scenario, with an entry rate into counseling equal to 20 percent. We also consider two versions of these policy experiments. In the first, the policy is permanent: the entry rate into counseling remains constant over time from time $t = 0$. In the second, it is transitory: some workers enter into counseling at time $t = 0$ only. Then, these workers remain counseled until they find a job and other workers cannot benefit from counseling.[14] As in the previous section, in all the simulations, the counseled have a comparative advantage which increases their relative probability of finding a job by 20 percent ($\delta = 1.2$).

5.4.1 Permanent policy

In the baseline scenario the entry rate into counseling, μ, is equal to 5 percent, which entails that 36 percent of the unemployed are counseled in steady state. In the 'light' scenario, where μ equals 1 percent, it turns out that 5.2 percent of the unemployed are counseled in steady state. In the 'intensive' scenario, with an entry rate into counseling equal to 20 percent, 69 percent of the unemployed are counseled in steady state. Figure 5.6 shows the dynamics of the share of counseled workers for the three cases.

[14]The simulations are made with Dynare, a collection of MATLAB routines which solve non-linear models with forward looking variables. Information about Dynare can be found in Juillard (1996) and at (http://www.cepremap.cnrs.fr/dynare/). The simulations make is necessary to write the model in discrete time; the discrete time version of the model is presented in appendix.

Figure 5.6: The evolution of the share of counseled workers (y-axis) over time (x-axis, by month) for $\mu = 0.01$ (continuous line), $\mu = 0.05$ (crosses) and $\mu = 0.2$ (circles).

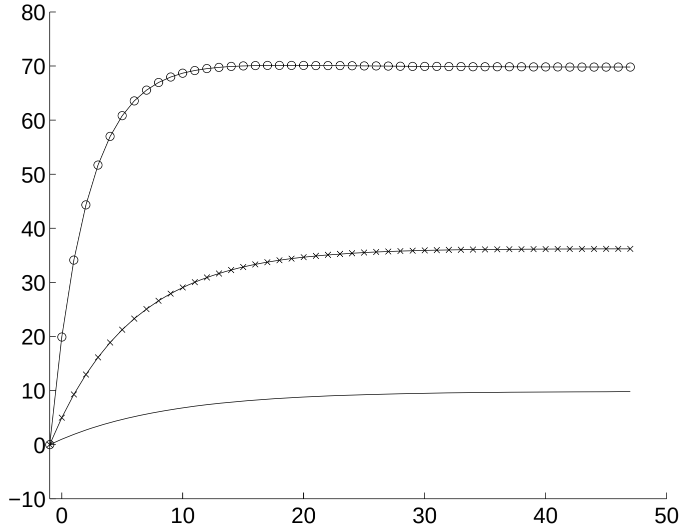


Figure 5.7: The evolution of the baseline arrival rate of job offers, $\theta q(\theta)$, (y-axis) over time (x-axis, by month) for $\mu = 0.01$ (continuous line), $\mu = 0.05$ (crosses) and $\mu = 0.2$ (circles).

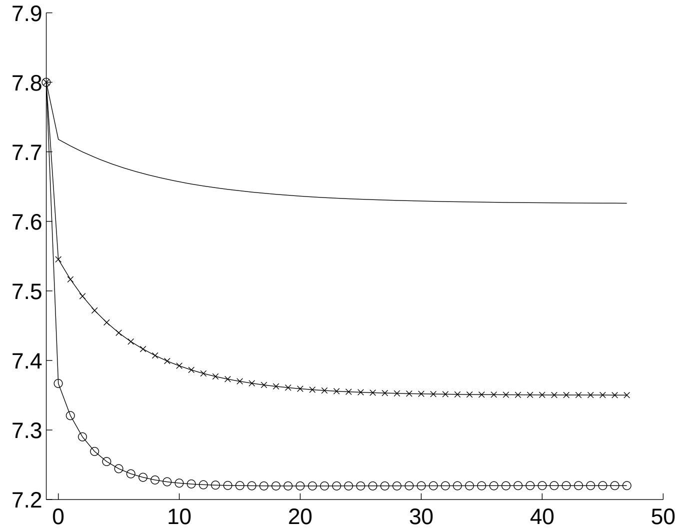


Figure 5.8: The evolution of the unemployment rate (y-axis) over time (x-axis, by month) for $\mu = 0.01$ (continuous line), $\mu = 0.05$ (crosses) and $\mu = 0.2$ (circles).

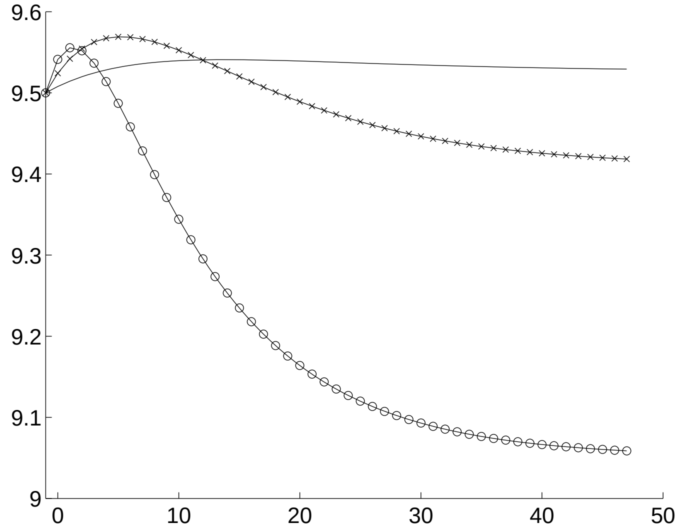


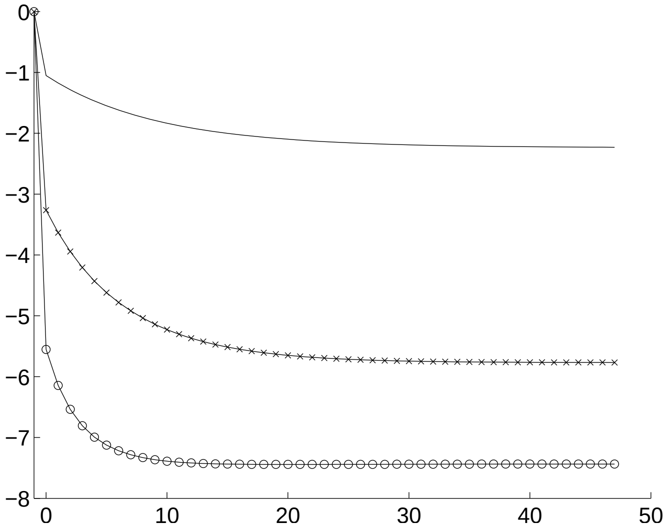
Figure 5.7 shows that the baseline arrival rate of job offers decreases monotonically with time. The baseline arrival rate of job offers adjusts more rapidly to its steady state value when the entry rate into counseling is bigger. However, in all cases considered here, the arrival rate of job offers is very close to its steady state value after one year.

Figure 5.8 shows that the unemployment rate dynamics are not always monotonic. There is an overshooting of the unemployment rate when the share of counseled job seekers is sufficiently large. This phenomenon is the consequence of the interaction between the positive impact of counseling on the entry rate into employment of counseled job seekers and the equilibrium effects, which reduce the entry rate into employment of the non-counseled. When the entry rate into counseling is large enough, the drop in the baseline arrival rate of job offers, induced by equilibrium effects, dominates at the beginning, which induces an increase in the unemployment rate. Then, as time elapses, there are more and more counseled workers whose exit rate out of unemployment is relatively high.

Figure 5.8 leads us to stress that it is important to account for the dynamics of the unemployment rate when evaluating the equilibrium effects of counseling. A priori, it could be possible to estimate the equilibrium effects of counseling by gathering data on similar employment pools in which there are different proportions of counseled individuals. However, this strategy can lead to very different conclusions according to the time horizon at which the evaluation is done. In the baseline scenario, where the entry rate into counseling amounts to 5 percent, the evaluation of the equilibrium effects 6 months after the introduction of the policy leads to the conclusion that they significantly increase unemployment. However, there are no significant effects on the unemployment rate beyond two years.

We also compute the dynamics of the evaluation error

Figure 5.9: The error (as a percentage of the naive evaluation δ) in the evaluation of the effect of counseling on the exit rate out of unemployment of counseled workers (y-axis) over time (x-axis, by month) for $\mu = 0.01$ (continuous line), $\mu = 0.05$ (crosses) and $\mu = 0.2$ (circles).

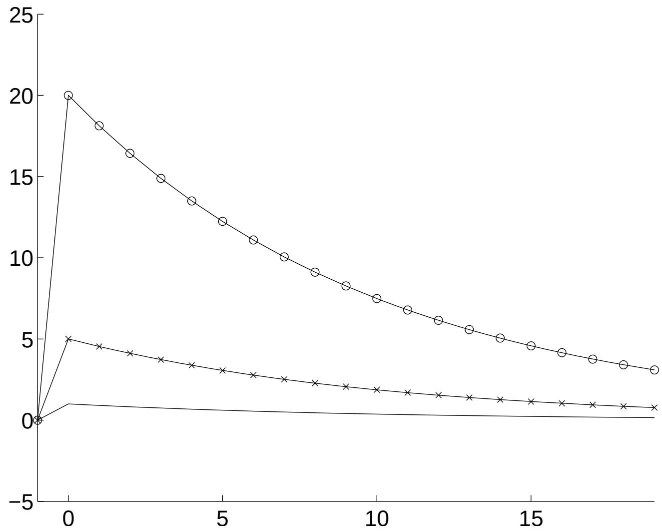


$\left[\theta_t q(\theta_t) - \theta_0 q(\theta_0)\right] / \theta_0 q(\theta_0)$. The true treatment effect on the treated is over-estimated by the naive evaluation. As shown by figure 5.7, the size of the error increases with time because it takes time to increase the number of counseled workers and then to get sizeable equilibrium effects.

5.4.2 Transitory policy

Now, we look at situations where some workers benefit from counseling at date zero and remain counseled until they find a job. The other job seekers, who do not benefit from counseling at date zero, are never counseled. Figure 5.10 displays the evolution of the share of counseled workers over time. Figure 5.11 shows the corresponding evolution of the unemployment rate. It turns out that counseling always decreases the unemployment rate, contrary to the case where the policy is permanent. The difference between the two cases comes from the role of the expectations of non-counseled workers. When the policy is permanent, non-counseled workers anticipate that they will benefit from counseling in the future. Therefore, their reservation wage and then their bargained wage increase (as shown in Figure 5.3 above). When the policy is transitory, non-counseled job seekers know that they will never benefit from counseling. Therefore, their reservation wage does not increase. Actually, their reservation wage decreases because the baseline arrival rate of job offers, $\theta q(\theta)$, drops when some workers are counseled, as shown by Figure 5.12. Note that there is a spike in the job offer arrival rate at the time of the policy shock. This is due to the assumption made in the discrete time version of the model presented in appendix. At date zero, there is no counseled worker ready to be hired since vacant

Figure 5.10: The evolution of the share of counseled workers (y-axis) over time (x-axis, by month) for $\mu = 0.01$ (continuous line), $\mu = 0.05$ (crosses) and $\mu = 0.2$ (circles).



jobs posted at date t are matched with workers unemployed at date $t-1$ (recall that unemployment is a predetermined variable). Moreover, at date zero, non-counseled job seekers reduce their reservation wage because they anticipate that the baseline arrival rate of job offers is going to decrease in the near future. The combination of these two phenomena increases the value of job vacancies, and then job creation at date zero. At date one, vacant jobs meet counseled job seekers whose reservation wage is higher. This is detrimental to job creation, as shown by Figure 5.12.

The comparison of the impact of transitory and permanent policies highlights the role of anticipations. When the policy is permanent, it turns out that a non negligible share of its impact on the unemployment rate is induced by the reaction of non-counseled job seekers. The rise in their reservation wage, and then in their bargained wage, induced by the expectation to benefit from counseling in the future dampens job creation. This phenomenon implies that permanent counseling increases unemployment when a small share of job seekers are counseled whereas counseling always decreases unemployment when it is transitory. Accordingly, it can be misleading to conclude that a truly successful transitory policy will remain successful when it becomes permanent.

5.5 Conclusion

Our paper stresses that it is worth accounting for equilibrium effects in the effort to provide a proper evaluation of counseling policies. Neglecting such effects could lead to the conclusion that counseling reduces steady state unemployment although its true effect could be the opposite. A striking result obtained in the paper is

Figure 5.11: The evolution of the unemployment rate (y-axis) over time (x-axis, by month) for $\mu = 0.01$ (continuous line), $\mu = 0.05$ (crosses) and $\mu = 0.2$ (circles).

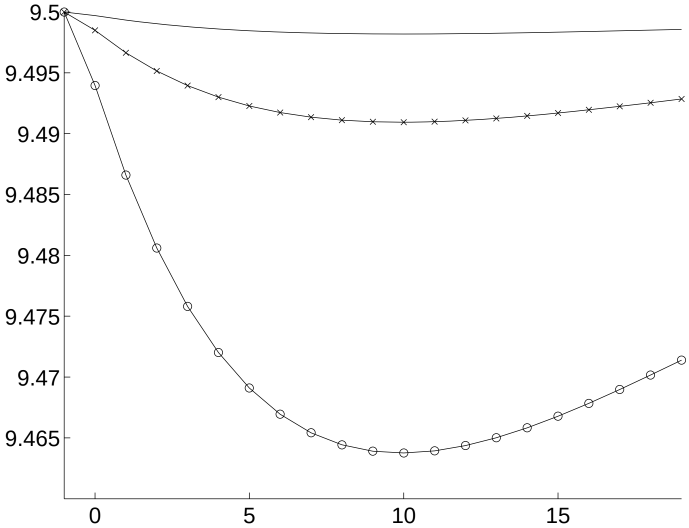
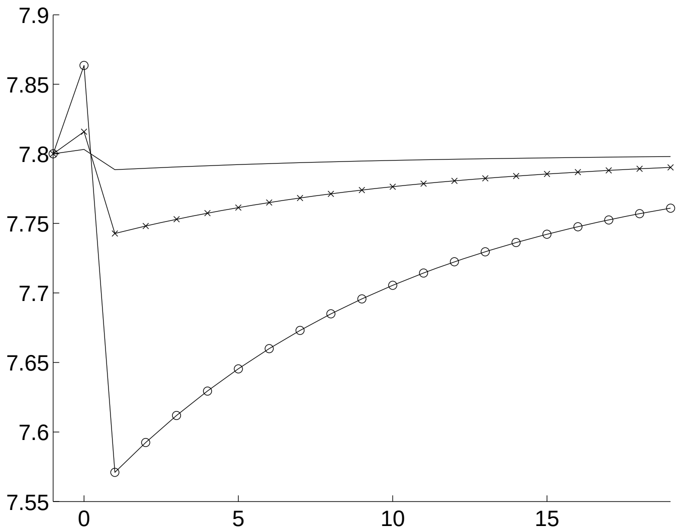


Figure 5.12: The evolution of the baseline arrival rate of job offers, $\theta q(\theta)$, (y-axis) over time (x-axis, by month) for $\mu = 0.01$ (continuous line), $\mu = 0.05$ (crosses) and $\mu = 0.2$ (circles).



that this type of error can arise when the size of the treatment group is small. It also turns out that it can be wrong to conclude that a truly successful transitory policy remains successful when it becomes permanent. This result is important to the extent that many policy evaluations rely on temporary experiments of policies. Typically, a policy is evaluated during a transitory period. Then, it is often assumed that this evaluation provides relevant information to evaluate the effect of the policy that will be implemented permanently. Our analysis shows that this is not always the case.

Appendix A: The model in discrete time

The aim of this appendix is to present the discrete time version of the continuous time model presented in the text. Unemployment rates are predetermined. During period t, matching involve the unemployed populations inherited from the previous period with the job vacancies posted in period t. To make clear that unemployment rates are predetermined, we index them by $t-1$. The timing of events within each period is the following: production takes place, firms post vacant jobs, jobs and unemployed workers are matched, jobs are destroyed at rate λ and, finally, payments are made. The assumptions about timing allow us to write the system of six equations from (5.11) to (5.16) that defines the equilibrium value of $(S_n, S_c, \theta, \Delta, u_n, u_c)$ as follows

$$\begin{aligned}
u_{n,t} &= \left[1-\mu-\theta_t q\left(\theta_t\right)\right] u_{n,t-1} + \lambda\left(1-u_{n,t-1}-u_{c,t-1}\right) \\
u_{c,t} &= \left[1-\delta\theta_t q\left(\theta_t\right)\right] u_{c,t-1} + \mu u_{n,t-1} \\
\text{where } \theta_t &= \frac{v_t}{\delta u_{c,t-1}+u_{n,t-1}}
\end{aligned}$$

$$\begin{aligned}
\frac{c}{q(\theta_t)(1-\beta)} &= \frac{u_{n,t-1}}{\delta u_{c,t-1}+u_{n,t-1}} S_{n,t+1} + \frac{\delta u_{c,t-1}}{\delta u_{c,t-1}+u_{n,t-1}} S_{c,t+1} \\
S_{n,t} &= \frac{1}{1+r}\left[y-z+\left[1-\lambda-\beta\theta_t q(\theta_t)\right] S_{n,t+1} - \mu\Delta_{t+1}\right] \\
S_{c,t} &= \frac{1}{1+r}\left[y-z+\left[1-\lambda-\delta\beta\theta_t q(\theta_t)\right] S_{c,t+1} - \lambda\Delta_{t+1}\right] \\
\Delta_t &= \frac{1}{1+r}\left[\beta\theta_t m(\theta_t)(\delta S_{c,t+1}-S_{n,t+1}) + (1-\mu)\Delta_{t+1}\right]
\end{aligned}$$

Because agents are risk-neutral per period social welfare can be written as the production net of the cost of vacant jobs. Thus, we define the period welfare as:

$$\omega_t = y(1-u_{n,t-1}-u_{c,t-1}) + z(u_{n,t-1}+u_{c,t-1}) - c\theta_t(u_{n,t-1}+\delta u_{c,t-1}).$$

The discounted present value of intertemporal welfare, Ω_t, is written:

$$\Omega_t = \beta\omega_t + \beta\Omega_{t+1}$$

Appendix B: The impact of counseling on labor market tightness

The aim of this appendix is to analyze the impact of changes in the entry rate into counseling, represented by parameter μ, on labor market tightness. We define $T = \theta q(\theta)$. For the sake of simplicity, we consider the special case where $q(\theta) = \theta^{-1/2}$ so that $T = 1/q(\theta)$.

We can write a system of 3 equations that define the 3 variables T, S_n, S_c in steady state. To obtain this system, we use equation (5.13) to substitute Δ into equations (5.11) and (5.12). We get:

$$\left[r + \lambda + \beta\delta T + \frac{\lambda}{r+\mu}\beta\delta T\right] S_c - \frac{\lambda}{r+\mu}\beta T S_n = y - z \tag{5.51}$$

$$\left[r + \lambda + \beta T - \frac{\mu}{r+\mu}\beta T\right] S_n + \frac{\mu}{r+\mu}\beta\delta T S_c = y - z. \tag{5.52}$$

From the free entry condition (5.14) and equations (5.15) and (5.16) we obtain the third equation:

$$\frac{c}{(1-\beta)}T = \frac{T}{\mu+T}S_n + \frac{\mu}{\mu+T}S_c. \tag{5.53}$$

Let us differentiate this system to find the sign of the derivative $\mathrm{d}T/\mathrm{d}\mu$. We can proceed by steps. We begin to differentiate the free entry condition (5.53):

$$\left[\frac{c}{(1-\beta)} - \frac{\mu}{(\mu+T)^2}(S_n - S_c)\right]\mathrm{d}T = -\frac{T}{(\mu+T)^2}(S_n - S_c)\,\mathrm{d}\mu + \frac{T}{\mu+T}\mathrm{d}S_n + \frac{\mu}{\mu+T}\mathrm{d}S_c \tag{5.54}$$

By using the free entry condition (5.53) again, it turns out that the factor before $\mathrm{d}T$ is positive. To go further in the analysis of changes in μ we need to differentiate equations (5.51) and (5.52) and solve for $\mathrm{d}S_n$ and $\mathrm{d}S_c$. Here is the solution written in compact terms:

$$A\mathrm{d}S_c = -B_c\mathrm{d}T + C_c\mathrm{d}\mu \tag{5.55}$$

$$A\mathrm{d}S_n = -B_n\mathrm{d}T - C_n\mathrm{d}\mu \tag{5.56}$$

where

$$\begin{aligned}
A &= \left(r+\lambda+\delta\beta T + \frac{\lambda}{r+\mu}\delta\beta T\right)\left(r+\lambda+\frac{r}{r+\mu}\beta T\right) + \frac{\lambda}{r+\mu}\beta T\frac{\mu}{r+\mu}\beta\delta T \\
B_c &= \left(r+\lambda+\frac{r}{r+\mu}\beta T\right)\left(\delta\beta S_c + \frac{\lambda}{r+\mu}\beta\left(\delta S_c - S_n\right)\right) + \frac{\lambda}{r+\mu}\beta T\left(\beta S_n + \frac{\mu}{r+\mu}\beta\left(\delta S_c - S_n\right)\right) \\
C_c &= \frac{\lambda}{(r+\mu)^2}\beta T\left(\delta S_c - S_n\right)(r+\lambda) \\
B_n &= \frac{\mu}{r+\mu}\beta\delta S_c(r+\lambda) + \beta S_n\frac{1}{r+\mu}\left(r\left(r+\lambda+\delta\beta T\right) + \lambda\delta\beta T\right) \\
C_n &= \beta T\left(\delta S_c - S_n\right)\frac{\lambda}{(r+\mu)^2}\beta\delta T + \beta T\left(\delta S_c - S_n\right)\frac{r}{(r+\mu)^2}\left(r+\mu+\beta\delta T\right)
\end{aligned}$$

Noticing that $(\delta S_c - S_n) > 0$, it appears that A, B_c, C_c, B_n, C_n are positive. Substituting the expressions of $\mathrm{d}S_c$ and $\mathrm{d}S_n$ defined by equations (5.55) and (5.56) into the free entry condition (5.54) we can see that the sign of $\mathrm{d}T/\mathrm{d}\mu$ is negative if

$$-\beta T\left(\delta S_c - S_n\right)\frac{1}{(r+\mu)^2}\left[\beta\delta T^2(r+\lambda) + r(r+\mu)T - \mu\lambda(r+\lambda)\right]$$

is also negative. The sign of this expression is a priori ambiguous. However, it is easy to check that it is negative when μ, the entry rate into counseling, is small.

Bibliography

Abbring, J. H., and G. J. van den Berg (2003): “The Nonparametric Identification of Treatment Effects in Duration Models,” *Econometrica*, 71(5), 1491–1517.

Acemoglu, D., and R. Shimer (2000): “Productivity gains from unemployement insurance,” *European Economic Review*, 44(7), 1195–1124. 6, 66

Addison, J. T., and M. L. Blackburn (2000): “The effects of unemployment insurance on postunemployment earnings,” *Labour Economics*, 7(1), 21–53. 6, 66, 68

Addison, J. T., and P. Portugal (2008): “How do different entitlements to unemployment benefits affect the transitions from unemployment into employment?,” *Economics Letters*, 101(3), 206–209. 6, 68

Albrecht, J., G. van den Berg, and S. Vroman (2009): “The Aggregate Labor Market Effects of the Swedish Knowledge Lift Program,” *Review of Economic Dynamics*, 12(1), 129–146. 11, 134

Angrist, J., G. Imbens, and D. Rubin (1996): “Identification of Causal Effects Using Instrumental Variables,” *Journal of the American Statistical Association*, 91, 444–455. 10, 129, 133

Anwar, S., and H. Fang (2006): “An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence,” *American Economic Review*, 96(1), 127–151. 16

Aslund, O., and O. Nordstrom Skans (2007): “Do anonymous job application procedures level the playing field?,” Working Paper Series 2007:31, IFAU - Institute for Labour Market Policy Evaluation. 4, 16

Behaghel, Crepon, and Gurgand (2012): “Private and Public Provision of Counseling to Job-Seekers : Evidence from a Large Controlled Experiment,” . 10, 11, 119

Belzil, C. (2001): “Unemployment insurance and subsequent job duration: job matching versus unobserved heterogeneity,” *Journal of Applied Econometrics*, 16(5), 619–636. 6, 68

Bennmarker, H., K. Carling, and B. Holmlund (2007): “Do Benefit Hikes Damage Job Finding? Evidence from Swedish Unemployment Insurance Reforms,” *LABOUR*, 21(1), 85–120. 6, 68

Bertrand, M., and S. Mullainathan (2004): “Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination,” *American Economic Review*, 94(4), 991–1013. 4, 14, 23

BLANCHARD, O., AND A. LANDIER (2002): "The Perverse Effects of Partial Labour Market Reform: fixed–Term Contracts in France," *Economic Journal*, 112(480), F214–F244. 1

BLANCHET, D., M. BARLET, AND T. L. BARBANCHON (2009): "Microsimulation et modèles dagents : une approche alternative pour lévaluation des politiques demploi," *Économie et Statistique*, 429(1), 51–76. 2

BLUNDELL, R., M. COSTA DIAS, C. MEGHIR, AND V. R. J. (2004): "Evaluating the employment impact of a mandatory job search program," *Journal of the European Economic Association*, 2, 569–606. 11, 134

BOONE, J., P. FREDRIKSSON, B. HOLMLUND, AND J. VAN OURS (2007): "Optimal Unemployment Insurance with Monitoring and Sanctions," *Economic Journal*, 117, 399–421. 134

BOONE, J., AND J. C. VAN OURS (2009): "Why is there a spike in the job finding rate at benefit exhaustion?," CEPR Discussion Papers 7525, C.E.P.R. Discussion Papers. 73, 96

BRODATY, T., B. CRÉPON, AND D. FOUGÈRE (2000): "Using Matching Estimators to Evaluate Alternative Youth Employment Programs: Evidence from France, 1986-1988," CEPR Discussion Papers 2604, C.E.P.R. Discussion Papers.

CAHUC, P., AND F. FONTAINE (2009): "On the efficiency of job search with social networks," *Journal of Public Economic Theory*, 11(3), 323–64. 136

CAHUC, P., AND F. POSTEL-VINAY (2002): "Temporary jobs, employment protection and labor market performance," *Labour Economics*, 9(1), 63–91. 1

CAHUC, P., F. POSTEL-VINAY, AND J. ROBIN (2006): "Wage bargaining with on-the-job search: A structural econometric model," *Econometrica*, 74(2), 323–64. 138

CALIENDO, M., K. TATSIRAMOS, AND A. UHLENDORFF (2009): "Benefit Duration, Unemployment Duration and Job Match Quality: A Regression-Discontinuity Approach," IZA Discussion Papers 4670, Institute for the Study of Labor (IZA). 6, 66, 68, 69

CALMFORS, L. (1994): "Active Labor Market Policy and Unemployment - A Framework for the Analysis of Crucial Design Features.," *OECD Economic Studies*, 22(1), 7–47. 11, 134

CARD, CHETTY, AND WEBER (2007): "Cash-on-Hand and Competing Models of Intertemporal Behavior: New Evidence from the Labor Market," *The Quarterly Journal of Economics*, 122(4), 1511–1560. 6, 66, 68, 69, 79, 82, 83

——— (2008): "The spike at benefit exhaustion : Leaving the unemployment system or starting a new job ?," *American Economic Review*, 97(2), 113–118. 119

Card, D., J. Kluve, and A. Weber (2010): "Active Labour Market Policy Evaluations: A Meta-Analysis," *Economic Journal*, 120(548), F452–F477. 11

Card, D., and P. B. Levine (2000): "Extended benefits and the duration of UI spells: evidence from the New Jersey extended benefit program," *Journal of Public Economics*, 78(1-2), 107–138. 5, 68

Carling, K., B. Holmlund, and A. Vejsiu (2001): "Do Benefit Cuts Boost Job Finding? Swedish Evidence from the 1990s," *Economic Journal*, 111(474), 766–90. 6, 68

Cavaco, S., D. Fougère, and J. Pouget (2005): "Estimating the Effect of a Retraining Program for Displaced Workers on Their Transition to Permanent Jobs," IZA Discussion Papers 1513, Institute for the Study of Labor (IZA).

——— (2009): "Estimating the Effect of a Retraining Program on the Re-Employment Rate of Displaced Workers," IZA Discussion Papers 4227, Institute for the Study of Labor (IZA).

Centeno, M. (2004): "The Match Quality Gains from Unemployment Insurance," *Journal of Human Resources*, 39(3). 6, 68

Centeno, M., and A. Novo (2009): "Reemployment wages and UI liquidity effect: a regression discontinuity approach," *Portuguese Economic Journal*, 8(1), 45–52. 6, 66, 69, 82

Chebira, A. (2005): "Líndispensable manuel orthographique des prénoms francais dórigine arabe et musulmane," Discussion paper, Ed APIC. 23

Crepon, Duflo, Gurgand, Rathelot, and Zamora (2012): "Do Labor Market Policies have Displacement Effect ? Evidence from a Clustered Randomized Experiment," . 10

Crépon, B., M. Dejemeppe, and M. Gurgand (2005): "Counseling the Unemployed: Does It Lower Unemployment Duration and Recurrence?," IZA Discussion Papers 1796, Institute for the Study of Labor (IZA). 11, 136

Crépon, B., and R. Desplatz (2001): "Une nouvelle évaluation des effets des allégements de charges sociales sur les bas salaires," *Économie et Statistique*, 348(1), 3–34. 2

Das, M., W. K. Newey, and F. Vella (2003): "Nonparametric Estimation of Sample Selection Models," *Review of Economic Studies*, 70(1), 33–58. 113

Davidson, C., and S. Woodbury (1993): "The Displacement Effect of Reemployment Bonus Programs.," *Journal of Labor Economics*, 11, 575–605. 11, 134

Dormont, B., D. Fougere, and A. Prieto (2001): "L'effet de l'allocation unique degressive sur la reprise d'emploi," *Economie et Statistique*, (343). 69

DUFLO, E., R. GLENNERSTER, AND M. KREMER (2008): *Using Randomization in Development Economics Research: A Toolkit*vol. 4 of *Handbook of Development Economics*, chap. 61, pp. 3895–3962. Elsevier. 7, 9

DUGUET, E., Y. L´ HORTY, D. MEURS, AND P. PETIT (2010): "Measuring Discriminations: an Introduction," *Annals of Economics and Statistics*, (99/100), 991–1013. 4, 14

FELOUZIS, G. (2003): "La ségrégation ethnique au collège et ses conséquences," *Revue Francaise de Sociologie*, 44(3), 413–447. 23

FISMAN, R., S. IYENGAR, E. KAMENICA, AND I. SIMONSON (2008): "Racial Preferences in Dating," *Review of Economic Studies*, 75, 117–132. 16

FOUGÈRE, D., J. PRADEL, AND M. ROGER (2009): "Does the public employment service affect search effort and outcomes?," *European Economic Review*, 53(7), 846–869. 11

FREMIGACCI, F. (2010): "Maximum Benefits Duration and Older Workers' Transitions out of Unemployment : a Regression Discontinuity Approach," CEE working Papers 10-12, CEE. 68

GAUTIER, MULLER, VAN DER KLAAUW, ROSHOLM, AND SVARER (2012): "Estimating Equilibrium Effect of Job Search Assistance," .

GAUTIER, P. A., AND B. VAN DER KLAAUW (2012): "Selection in a field experiment with voluntary participation," *Journal of Applied Econometrics*, 27(1), 63–84.

GOLDIN, C., AND C. ROUSE (2000): "Orchestrating impartiality : the impact of " Blind " auditions on female musicians," *American Economic Review*, 90(4). 4, 15, 16

HAHN, J., P. TODD, AND W. V. D. KLAAUW (2001): "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 69(1), pp. 201–209. 74

HAM, J. C., AND R. J. LALONDE (1996): "The Effect of Sample Selection and Initial Conditions in Duration Models: Evidence from Experimental Data on Training," *Econometrica*, 64(1), 175–205. 80

HECKMAN, J., L. LOCHNER, AND C. TABER (1998): "General Equilibrium Treatment Effects: A Study of Tuition Policy.," *American Economic Review*, 82(2), 381–386. 12, 134

HECKMAN, J. J. (1976): "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models," *The Annals of Economic and Social Measurement*, 5, 475–492.

——— (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47(1), 153–61.

HECKMAN, J. J., AND J. A. SMITH (1998): "Evaluating the Welfare State," NBER Working Papers 6542, National Bureau of Economic Research, Inc.

HOROWITZ, J., AND C. MANSKI (2000): "Nonparametric Analysis of Randomized Experiments with Missing Covariate and Outcome Data," *Journal of the American Statistical Association*, 95. 110

HOROWITZ, J. L., AND C. F. MANSKI (1995): "Identification and Robustness with Contaminated and Corrupted Data," *Econometrica*, 63(2), 281–302.

HOROWITZ, J. L., AND C. F. MANSKI (1998): "Censoring of outcomes and regressors due to survey nonresponse: Identification and estimation using weights and imputations," *Journal of Econometrics*, 84(1), 37–58.

HUNT, J. (1995): "The Effect of Unemployment Compensation on Unemployment Duration in Germany," *Journal of Labor Economics*, 13(1), 88–120. 6, 68

IMBENS, G. W., AND T. LEMIEUX (2008): "Regression discontinuity designs: A guide to practice," *Journal of Econometrics*, 142(2), 615–635. 74

JUILLARD, M. (1996): "Dynare : a program for the resolution and simulation of dynamic models with forward variables through the use of a relaxation algorithm," CEPREMAP Working Papers (Couverture Orange) 9602, CEPREMAP. 147

JURAJDA, S. (2002): "Estimating the effect of unemployment insurance compensation on the labor market histories of displaced workers," *Journal of Econometrics*, 108(2), 227–252. 6, 68

KLUVE, J. (2010): "The effectiveness of European active labor market programs," *Labour Economics*, 17(6), 904–918. 11

KRAUSE, A., U. RINNE, AND K. F. ZIMMERMANN (2011): "Anonymous Job Applications of Fresh Ph.D. Economists," IZA Discussion Papers 6100, Institute for the Study of Labor (IZA). 5

KREMER, M., E. MIGUEL, AND R. THORNTON (2009): "Incentives to Learn," *The Review of Economics and Statistics*, 91(3), 437–456. 111

KRUEGER, A. B. (1999): "Experimental Estimates Of Education Production Functions," *The Quarterly Journal of Economics*, 114(2), 497–532.

KYYRÄ, T., AND V. OLLIKAINEN (2008): "To search or not to search? The effects of UI benefit extension for the older unemployed," *Journal of Public Economics*, 92(10-11), 2048–2070. 6, 68

Lalive, R. (2007): "Unemployment Benefits, Unemployment Duration, and Post-Unemployment Jobs: A Regression Discontinuity Approach," *American Economic Review*, 97(2), 108–112. 6, 66, 69

——— (2008): "How do extended benefits affect unemployment duration A regression discontinuity approach," *Journal of Econometrics*, 142(2), 785–806. 6, 68

Lalive, R., J. V. Ours, and J. ZweimÃ$\frac{1}{4}$ller (2006): "How Changes in Financial Incentives Affect the Duration of Unemployment," *Review of Economic Studies*, 73(4), 1009–1038. 6, 68

Lalive, R., and J. Zweimuller (2004): "Benefit entitlement and unemployment duration: The role of policy endogeneity," *Journal of Public Economics*, 88(12), 2587–2616. 6, 68

LaLonde, R. J. (1986): "Evaluating the Econometric Evaluations of Training Programs with Experimental Data," *American Economic Review*, 76(4), 604–20. 7, 110

Lee, D. (2009a): "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *Review of Economic Studies*, 76, 1071–1102.

Lee, D. S. (2009b): "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *Review of Economic Studies*, 76(3), 1071–1102. 9, 81, 110, 111

Lise, J., S. Seitz, and J. Smith (2005): "Equilibrium Policy Experiments and the Evaluation of Social Programs," . 11, 134

Malcomson, J. (2011): *Individual Employment Contracts*vol. 3B of *Handbook of Labor Economics*, chap. 35, pp. 2291–2372. Elsevier. 138

Manski, C. (1989): "Anatomy of the Selection Problem," *Journal of Human Resources*, 24(3), 343–60. 110

Marimon, R., and F. Zilibotti (1997): "Unemployment vs mismatch of talent : reconsidering unemployment benefits," *NBER Working Papers*, (6038). 6, 66

Masters, A. (2000): "Retraining the unemployed in a model of equilibrium employment," *Bulletin of Economic Research*, 52, 323–340. 135

McCrary, J. (2008): "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, 142(2), 698–714. 75

Meghir, C. (2006): "Dynamic models for policy evaluation," IFS Working Papers W06/08, Institute for Fiscal Studies. 10, 134

MEYER, B. D. (1990): "Unemployment Insurance and Unemployment Spells," *Econometrica*, 58(4), 757–82. 5, 68

NEWEY, W. K., AND D. MCFADDEN (1986): "Large sample estimation and hypothesis testing," in *Handbook of Econometrics*, ed. by R. F. Engle, and D. McFadden, vol. 4 of *Handbook of Econometrics*, chap. 36, pp. 2111–2245. Elsevier. 128

ORES (2007): "Porter un prénom arabe ou musulman est-il discriminant dans lénseignement supérieur ? Orientation et poursuite détudes dans lénseignement supérieur," Discussion paper, Pole universitaire Lille - Nord - Pas-de-Calais.

PISSARIDES, C. (1979): "Job Matchings with State Employment Agencies and Random Search," *Economic Journal*, 89, 818–33. 136

PISSARIDES, C. A. (2000): *Equilibrium Unemployment Theory, 2nd Edition*, vol. 1 of *MIT Press Books*. The MIT Press.

PRICE, J., AND J. WOLFERS (2010): "Racial Discrimination Among NBA Referees," *The Quarterly Journal of Economics*, 125(4), 1859–1887. 16

PUHANI, P. A. (2000a): "The Heckman Correction for Sample Selection and Its Critique," *Journal of Economic Surveys*, 14(1), 53–68.

PUHANI, P. A. (2000b): "Poland on the dole: The effect of reducing the unemployment benefit entitlement period during transition," *Journal of Population Economics*, 13(1), 35–44. 6, 68

RIDDER, G., AND J. VIKSTROM (2011): "Bounds on treatment effects on transitions," Working Paper Series 6, IFAU - Institute for Labour Market Policy Evaluation. 79

RUBIN, D. (1978): "Baysian Inference for Causal Effects; the Role of Randomization," *The Annals of Statistics*, 6, 34–58. 7, 10, 133

SCHMIEDER, J. F., T. VON WACHTER, AND S. BENDER (2012a): "The Effects of Extended Unemployment Insurance over the Business Cycle: Evidence from Regression Discontinuity Estimates over 20 Years," *Quarterly Journal of Economics*. 68

SCHMIEDER, J. F., T. M. VON WACHTER, AND S. BENDER (2012b): "The Long-Term Effects of Unemployment Insurance Extensions on Employment," NBER Working Papers 17814, National Bureau of Economic Research, Inc. 69

SIMON, P. (1998): "Nationalité et origine dans la statistique francaise," *Population*, 53(3), 541–567. 24

TATSIRAMOS, K. (2009): "Unemployment Insurance in Europe: Unemployment Duration and Subsequent Employment Stability," *Journal of the European Economic Association*, 7(6), 1225–1260. 6, 68

Uusitalo, R., and J. Verho (2010): "The effect of unemployment benefits on re-employment rates: Evidence from the Finnish unemployment insurance reform," *Labour Economics*, 17(4), 643–654. 6, 68

Van den Berg, G., and B. van der Klaauw (2006): "Counseling and monitoring of unemployed workers: theory and evidence from a controlled social experiment," *International Economic Review*, 47, 895–936. 136

van den Berg, G. J. (1990): "Nonstationarity in Job Search Theory," *Review of Economic Studies*, 57(2), 255–77. 67

Van der Linden, B. (2005): "Equilibrium Evaluation of Active Labor Market Programmes Enhancing Matching Effectiveness," IZA Discussion Papers 1526, Institute for the Study of Labor (IZA). 11, 12, 134

van Ours, J. C., and M. Vodopivec (2006): "How Shortening the Potential Duration of Unemployment Benefits Affects the Duration of Unemployment: Evidence from a Natural Experiment," *Journal of Labor Economics*, 24(2), 351–350. 6, 68

——— (2008): "Does reducing unemployment insurance generosity reduce job match quality?," *Journal of Public Economics*, 92, 684–695. 6, 66, 68

Vytlacil, E. (2002): "Independence, Monotonicity, and Latent Index Models: An Equivalence Result," *Econometrica*, 70(1), 331–341. 116

Winter-Ebmer, R. (1998): "Potential Unemployment Benefit Duration and Spell Length: Lessons from a Quasi-Experiment in Austria," *Oxford Bulletin of Economics and Statistics*, 60(1), 33–45. 6, 68

Evaluation de politiques publiques sur le marché du travail

Résumé : L'objectif de cette thèse est d'apporter un éclairage sur l'efficacité des politiques publiques sur le marché du travail.

Dans les deux premiers chapitres, nous proposons des évaluations empiriques de deux politiques publiques sur le marché du travail français:

- le CV anonyme: le bloc état-civil est supprimé du CV (premier chapitre). L'anonymisation réduit les écarts d'accès aux entretiens entre les femmes et les hommes. Cet effet limité aux offres d'emploi pour lequelles à la fois des hommes et des femmes postulent semble se prolonger aux phases ultérieures du recrutement, jusqu'à l'embauche;

- générosité de l'assurance chômage : une augmentation de 8 mois de durée maximale d'assurance chômage n'affecte pas la qualité de l'emploi trouvé, même si elle ralentit le retour à l'emploi (second chapitre).

Ces évaluations s'appuient sur des données expérimentales ou quasi expérimentales permettant l'identification de causalité. A cet effet, nous comparons un groupe traité et un groupe de contrôle qui sont statistiquement identiques avant la mise en place du traitement. Ce cadre d'analyse n'est pour autant pas exempt de difficultés méthodologiques. Dans les deux derniers chapitres, nous abordons deux difficultés méthodologiques des évaluations microeconométriques des programmes d'assistance aux chômeurs:

- comparabilité ex post des groupes traité et témoin en présence d'attrition différenciée entre groupes expérimentaux (troisième chapitre)

- effets d'équilibre affectant le groupe témoin (quatrième chapitre)

Mots clés : Evaluation; Discrimination; CV anonyme; Assurance chômage; Accompagnement des chômeurs; modèle frictionnel d'appariement; Régression discontinue; Evaluation randomisée; non réponse

Essays on Labor Market Policies Evaluation

Abstract: In the first two chapters, we estimate the impact of two labor market policies in the French context:

- anonymous applications: information, such as name, gender, age, nationality and address, is erased from resumes sent to employers (first chapter). Anonymous applications limit differential treatments based on gender and counter homophily.
- Unemployment Insurance generosity: job losers receive benefits to prevent revenue loss and to subsidize job search (second chapter). We show that an increase of 8 months in potential benefit duration does not affect match quality, while it slows down unemployment exits to jobs.

We devote special attention to identify causal impacts and thus rely on experimental or quasi experimental evidence to perform our empirical evaluation exercises. Namely, we compare treated and control groups which are ex ante statistically identical. However, we discuss two potential caveats of microeconometric evaluations in the context of Job Search Assistance evaluation, one ”practical” in the third chapter and one ”theoretical” in the fourth chapter:

- ex post comparability of control and treated groups when sample attrition can be different among experimental groups (third chapter);
- uncontrolled contamination between control and treated group through market interaction (fourth chapter).

Keywords: Labor Market Policy evaluation; Discrimination; Anonymous application; Unemployment Insurance; Job Search Assistance; Search and Matching model; Regression Discontinuity; Randomized evaluation; Survey non response
JEL Codes: J64 J65 J68 C41 C31 C93 J6 J71 J78