THE
ECONOMIC
JOURNAL

The Economic Journal, 1–41 https://doi.org/10.1093/ej/ueae095
Advance Access Publication Date: 5 November 2024

EMPLOYER SCREENING AND OPTIMAL UNEMPLOYMENT INSURANCE*

Mario Meier and Tim Obermeier

Field experiments show that employers are less likely to consider long-term unemployed job seekers for interviews. We study the implications for optimal unemployment insurance. Based on a structural model of job search and recruitment, estimated with German data, we analyse the optimal two-tier unemployment system. We find that screening makes the optimal initial benefit level 4 percentage points higher and the potential benefit duration seven months longer. Using an extended Baily–Chetty formula, we study the mechanisms through which screening affects the consumption smoothing gain and moral hazard cost of providing unemployment insurance and highlight the role of the externality from endogenous firm behaviour.

Recent evidence from field experiments indicates that employers take unemployment duration into account when making decisions about which applicants to invite for interviews. Kroft *et al.* (2013) found that the probability of being invited for an interview falls by 50% within the first six months of unemployment and argued that this can best be rationalised as screening behaviour. When job seekers differ in their productivity, less employable individuals are more likely to experience long unemployment spells, which can make firms reluctant to interview the long-term unemployed. Such statistical discrimination against the long-term unemployed can have severe consequences for job seekers, as a long unemployment spell can make it difficult to be considered for jobs even when the worker is qualified.

How should the optimal unemployment insurance (UI) system account for these 'stigma' effects? The presence of screening affects the optimal policy problem in various ways. The main goal of UI is to provide insurance while maintaining search incentives. The presence of screening leads to declining re-employment prospects the longer an individual is unemployed ('duration dependence'), which changes how strongly individuals value insurance, as well as

* Corresponding author: Tim Obermeier, School of Economics, University of Leicester, 266 London Road, Leicester LE2 1RQ, UK. Email: tim.obermeier@leicester.ac.uk

This paper was received on 22 December 2020 and accepted on 29 October 2024. The Editor was Barbara Petrongolo.

The data and codes for this paper are available on the Journal repository. They were checked for their ability to reproduce the results presented in the paper. The authors were granted an exemption to publish parts of their data because access to these data is restricted. However, the authors provided a simulated or synthetic dataset that allowed the Journal to run their codes. The synthetic/simulated data and the codes for the parts subject to exemption are also available on the Journal repository. They were checked for their ability to generate all tables and figures in the paper; however, the synthetic/simulated data are not designed to reproduce the same results. The replication package for this paper is available at the following address: https://doi.org/10.5281/zenodo.13985924.

We are especially grateful to Eckhard Janeba, Andreas Peichl and Michèle Tertilt for their invaluable advice and support with this project. Barbara Petrongolo (the editor) and two anonymous refeeres also provided key comments that helped us with the paper. In addition, we want to thank Jaap Abbring, Bruno Decreuse, Piotr Denderski, Georg Dürnecker, Javier Fernández-Blanco, Sebastian Findeisen, Ulrich Glogowsky, Andreas Haufler, Christian Holzner, Gregor Jarosch, Tom Krebs, Robert Kollmann, Kevin Lang, Tim Lee, Etienne Lehmann, Attila Lindner, Daniele Paserman, Fabien Postel-Vinay, Johannes Schmieder, Andrea Weber and Josef Zweimüller, as well as seminar participants at the University of Mannheim, University of Leicester, LMU Munich, IFS and TU Dortmund, and participants at many conferences and workshops. Financial support by the German Research Foundation (DFG) through CRC TR 224 (Project A03) is gratefully acknowledged. Finally, we greatly benefited from support by the state of Baden-Wuerttemberg through BwHPC and are grateful for their excellent technical support. This paper represents the authors' personal opinions and does not necessarily reflect the views of the Deutsche Bundesbank, the Eurosystem or their staff.

their incentives to search for jobs, and has implications for optimal UI for these reasons. In addition, the extent of screening itself depends on labour market conditions and the selection of different types of unemployed workers over the unemployment spell. For example, when UI programmes change their search incentives in a way that retains the most productive workers in the unemployment pool for longer, unemployment duration becomes a less informative signal and firms will discriminate less against the long-term unemployed. As a result, UI programmes can have an indirect effect on job-finding rates through their effect on firms' interviewing and hiring decisions, which the optimal UI system should also take into account.

Owing to these different considerations, it is not clear how the presence of screening affects the optimal UI system. To clarify this issue, we build a quantitative model of the job search and recruitment process and use the estimated model for policy analysis. In our model, workers decide on their search effort and savings. Firms receive multiple applications from workers and do not observe the productivity of each worker, but only their unemployment duration and a signal about productivity. Firms infer productivity by combining the information contained in unemployment duration and the signal and rank applicants by their expected productivity. Workers who are higher in the ranking have a higher probability of being hired. Taken together, the job-finding rate depends on both the search effort of workers and the *offer probability* (that is, the probability of receiving an offer conditional on having sent an application to a firm). The UI benefit system affects both of these aspects. Workers' search incentives depend on the generosity of the benefits they receive. As workers change their search behaviour, firms' beliefs about the productivity of the unemployed can change, as well as the pool of applicants. Firms can therefore adjust their interviewing and hiring choices, which leads to a change in the offer probabilities.

We estimate our model using German administrative data on job-finding rates and survey data on search effort, vacancies, applications and savings, including a comprehensive survey (the German Job Vacancy Survey) that contains information about the recruitment process from the perspective of firms. Our estimated model matches several empirical features of the job search and hiring process, namely, the decline in job-finding rates, the applications-per-vacancy ratio, the decline in interview rates and the decline in the search effort of agents. Using our estimated model, we then study optimal two-tier unemployment systems. These are schedules that are described by three policy variables: first, an initial benefit level (b_1); second, the number of months for which this benefit level is paid (the *potential benefit duration* (PBD), denoted D); and third, the second benefit level for the long-term unemployed (b_2). These two-tier systems reflect the UI policies that are in place in many countries.

Our analysis leads to three main messages. First, we find that screening has a significant impact on the optimal schedule. We compare the estimated screening model with a full-information benchmark, where the productivity of each applicant is perfectly revealed to firms so that they do not engage in statistical discrimination against the long-term unemployed. In the screening model, job-finding rates for the short-term unemployed are substantially higher than in the full-information benchmark, while the rates are lower for the long-term unemployed. Overall, the optimal initial benefit level (b_1) is 4 percentage points higher in the screening model (63% rather than 59%) than it would be in the absence of screening. In addition, the optimal potential benefit duration (D) increases substantially by seven months (twenty-six months compared with nineteen months). Therefore, we find that statistical discrimination, which has been documented by the audit study literature (e.g., Kroft *et al.*, 2013), has an important impact on optimal UI design.

Second, we study the mechanisms through which screening affects optimal policy design. Lehr (2017) and Kolsrud *et al.* (2018) showed that optimal policy in employer screening models can be characterised through an adjusted Baily–Chetty formula. The 'sufficient statistics' express the welfare gain from raising benefits as the difference between the consumption smoothing gain (the value of transferring resources from the employed to the unemployed), the moral hazard cost (how important revenue losses due to behavioural responses are relative to the mechanical revenue loss from raising benefits) and an externality term (capturing that changes in the UI policy can change offer probabilities). Using the framework from Kolsrud *et al.* (2018), we show that our result that the optimal initial benefit level (b_1) is higher in the screening model is driven by the combination of the following mechanisms: when employers screen applicants by duration, offer probabilities rise for short-term unemployed workers and fall for the long-term unemployed relative to the full-information benchmark. Workers search more intensely in the beginning of the spell to avoid long-term unemployment. Because of both of these effects, workers are more likely to find a job in the beginning of their unemployment spell. Therefore, there are fewer workers who receive b_1 and raising b_1 becomes less costly for the government, both in terms of the mechanical revenue loss and the behavioural response. The mechanical revenue loss falls more strongly and this effect increases the moral hazard cost (which is the ratio of the behavioural response and the mechanical revenue loss). Another mechanism that increases moral hazard costs is that workers internalise screening and their search effort becomes more responsive to benefits (the elasticity of effort with respect to benefits rises). The consumption smoothing gain goes up, as the wedge in the marginal utility of consumption between the unemployed and the employed grows. One of the key points of our analysis is that the planner also takes into account the fact that raising benefits affects firms' offer probabilities. The main mechanism for this effect is that the benefit system changes the composition of the applications coming from the long-term unemployed: when benefits are more generous, there are more high-type workers among long-term unemployed applicants. Firms' beliefs about the long-term unemployed improve and they are more likely to interview and hire them. As raising benefits increases offer rates, this leads to a positive externality effect on welfare. Furthermore, the impact of benefits on offer probabilities also lowers the elasticity of unemployment with respect to benefits, which leads to a fall in the moral hazard cost and offsets the earlier two effects that increased the moral hazard cost. Altogether, the consumption smoothing gain is 6.8% higher in the screening model than in the full-information benchmark, the moral hazard cost is 14.1% lower and the externality-adjusted moral hazard cost is 49.7% lower. Since the consumption smoothing gain from raising the initial benefit is now larger than the adjusted moral hazard cost, it is optimal for the planner to increase benefits. The reasoning for why the PBD is higher in the screening model is very similar. Overall, the hiring externality, which is new in employer screening models, plays an important role for our optimal policy results.

Third, we conduct a welfare analysis of who gains and loses when the government takes screening into account. Implementing the optimal schedule from the screening model—relative to wrongly imposing the optimal schedule from the full-information benchmark—leads to important distributional welfare effects. The lowest-type workers gain up to 1,777€, while the highest-type workers lose welfare equivalent to 937€. We get similar welfare gains and losses when re-estimating the full-information benchmark so that it fits the same data moments as the screening model. This corresponds to the case where the researcher 'wrongly' estimates the full-information model, while the labour market is better described by the screening model. In this case, low-type workers gain 1,258€, while high-type workers lose 730€. Taken together, these results suggest that taking screening into account—rather than basing the optimal policy on the full-information model—reduces inequality between low- and high-type workers.

Our paper contributes to the literature on (optimal) unemployment insurance by studying the role of employer screening for optimal UI policy design. Various papers have studied optimal UI policy design in partial equilibrium models (see, e.g., Baily, 1978; Gruber, 1997; Hopenhayn and Nicolini, 1997; Chetty, 2006; 2008; Shimer and Werning, 2006; 2008; Lentz, 2009; Pavoni, 2009) or, more recently, in general equilibrium settings with endogenous wages or market tightness (Michaillat, 2012; Lalive *et al.,* 2015; Marinescu, 2017; Landais *et al.*, 2018a,b; Hagedorn *et al.*, 2019). Even though this is a large literature, little attention has been paid to the role of employer screening. Two notable exceptions that are most closely related to our paper are Lehr (2017) and Kolsrud *et al.* (2018). Kolsrud *et al.* (2018) studied the optimal timing of UI benefits in the context of a Swedish reform. They extended the classic Baily–Chetty formula to the case of time-varying UI benefits in a partial equilibrium model and estimated the components of this formula (the 'sufficient statistics') based on the Swedish data. To discuss the robustness of their implementation, they also showed that the presence of employer screening would lead to an additional 'hiring externality' in the sufficient statistics formula, which however cannot be estimated in the Swedish data. Lehr (2017) also derived a sufficient statistics formula that accounts for employer screening by also including an externality term. Our paper extends and complements this work by studying the role of employer screening from the perspective of a structural model. In particular, we show how the introduction of screening affects optimal policy compared to the full-information benchmark and study the underlying mechanisms in terms of workers' search incentives and consumption, as well as firms' hiring behaviour. Furthermore, our analysis also suggests that the hiring externality, which has not previously been quantified, can be quantitatively important. More generally, this highlights the value of complementing reduced-form estimates of the sufficient statistics with model-based assessments of the hiring externality, in order to better assess the welfare gains from changing the UI policy.

In addition, our paper also contributes to a recent macroeconomic literature on the impact of employer screening in search and matching models of the labour market.[1] The novelty, relative to this literature, is studying the role of the UI policy in a screening model. Most closely related to our paper, Jarosch and Pilossoph (2019) used a random search framework to investigate the quantitative link between the decline in callback rates and duration dependence. While our broad modelling approach is similar to Jarosch and Pilossoph (2019), we add a combination of features, which is absent in their paper, but which is crucial for the analysis of the UI policy; in particular, endogenous search effort, risk aversion and savings. Fernandez-Blanco and Preugschat (2018) took a different approach by including screening in a directed search model that can account for endogenous wages. They showed that there is an equilibrium in which the labour market is not segmented, so that different types of workers apply for the same vacancies. Doppelt (2016) also built a directed search model with screening, focusing on the information contained in the labour market history of a worker's life cycle. Feng *et al.* (2019) also studied screening in a directed search model, characterised the types of equilibria and performed some illustrative policy experiments. Conceptually, an important difference is that the directed search models typically allow for endogenous wages. Wages are fixed in both Jarosch and Pilossoph (2019) and our model, as endogenous wages are more difficult to incorporate in random search models

[1] Early contributions about the potential role of employer screening were Lockwood (1991) and Kollmann (1994), who studied the informational content of unemployment duration. In addition, ranking models—where firms receive multiple applications and hire only one applicant—have a tradition going back to Blanchard and Diamond (1994), who assumed that firms hire the applicant with the shortest unemployment duration.

with multiple applicants per vacancy. While fixed wages are an important simplification relative to directed search models, this allows us to enrich our model in several other dimensions that are crucial for the analysis of the UI policy (in particular endogenous search effort and savings) and would be much more difficult to incorporate in a directed search model in the spirit of Fernandez-Blanco and Preugschat (2018).

The rest of the paper is organised as follows. In Section 1, we focus on the data and some descriptive facts. Section 2 presents the model and policy problem. Section 3 describes the estimation and discusses estimation results and model fit. In Section 4, we discuss welfare and the corresponding policy results and Section 5 concludes.

1. Data and Descriptive Facts

This section presents the data we use and empirical facts about job search behaviour and the hiring process.

1.1. *Data*

For the purpose of our analysis, we use data from Germany, as there are various datasets available that cover different aspects of the job search and hiring process. In Germany, most unemployed individuals receive unemployment benefits for up to twelve months of unemployment and are eligible for unemployment assistance if they remain unemployed for longer than twelve months. Older individuals are eligible for longer unemployment insurance payments, but we restrict our study to individuals who receive twelve months of benefits. Unemployed individuals receive benefits that amount to 60% or 67% of their past wage, depending on their marital status. After individuals run out of UI, they receive means-tested unemployment assistance benefits, which are around 40% for individuals who earn the average wage in our sample. Unemployment benefits are financed by social security contributions of workers and firms.[2]

The German setting allows us to base the design and estimation of our model on several datasets that contain information on job-finding rates, search effort and vacancies. First, we use the German social insurance data (IEB/SIAB), which provide information on the characteristics of the unemployed, in particular the length of their unemployment spell and their wage history. The data contain all individuals who have been unemployed at some time or regularly employed through an employment relationship that is subject to social insurance. We have access to a 2% random sample of the population and restrict ourselves to unemployment spells starting from 2000 until 2011. Second, we use the IZA Evaluation Dataset (IZA ED), which is a representative survey performed among UI entrants between June 2007 and May 2008. These data come from a panel where participants were interviewed up to four times after their period of unemployment commenced. The first interview took place close to the start of their unemployment spell. Additional interviews took place six, twelve and thirty-six months after they started receiving UI. Participants in this study were asked about their individual search effort (for example, the number of applications or number of search channels) and to report their reservation wage (that

[2] The German unemployment insurance system compares relatively well to unemployment insurance schemes in other developed countries, such as the United States and many other European countries. However, the system in the United States has a less generous potential benefit duration and replacement rate than Germany and no unemployment assistance system. For further details on the institutions in Germany, we refer the reader to Online Appendix C.

Table 1. *Descriptive Statistics.*

Variables	*N*	Mean	SD
Panel A: employment register			
Re-employment wage (euros)	55,362	1,605.33	(1,059.92)
Unemployment duration (months)	59,751	12.59	(12.71)
Female	59,751	0.445	(0.497)
Age	59,751	30.8	(9.12)
Married	59,751	0.324	(0.468)
Children	59,751	0.302	(0.459)
College	56,694	0.096	(0.294)
Apprenticeship	56,694	0.752	(0.432)
Panel B: IZA Evaluation Dataset			
Number of applications, month 1	6,815	13.49	(14.95)
Number of applications, month 6	377	9.15	(10.09)
Number of applications, month 12	1,710	8.11	(9.78)
Search channels, month 1	6,898	4.78	(1.8)
Panel C: Panel on Household Finances (quantiles)			
Net liquid assets (euros, p10)	295	−1,003	–
Net liquid assets (euros, p25)	295	−6	–
Net liquid assets (euros, p50)	295	247	–
Net liquid assets (euros, p75)	295	4,997	–
Net liquid assets (euros, p90)	295	40,497	–
Net assets (euros, including home, p50)	295	894	–
Panel D: Job Vacancy Survey			
Number of applicants	62,904	14.79	(36.96)
Time vacancy is open (days)	76,240	56.75	(67.08)

Notes: This table shows descriptive statistics from our different data sources. Panel A shows descriptive statistics from the administrative employment registers of individuals who experience their first unemployment spell at the time it starts. Panel B summarises search effort measures from the IZA Evaluation Dataset. Panel C uses the Bundesbank Panel on Household Finances for information on assets. In panel D, statistics from the IAB Job Vacancy Survey are shown. Here, *N* denotes the number of observations behind each statistic.

is, the minimum acceptable wage for them to exit unemployment and start a job). Third, we use the IAB Job Vacancy Survey (JVS), which is a representative survey conducted among firms on open vacancies and hiring decisions made by firms. The survey contains information on whether unemployed applicants were hired and how many applicants firms invite to an interview. Fourth, we use the Bundesbank Panel on Household Finances (PHF), which contains information on savings, liquid assets and debt levels. In these data, individuals are also asked to report whether they are employed or unemployed.

Table 1 summarises some of the main characteristics of the data sources. The average monthly re-employment wage after unemployment for job seekers is 1,605 euros. The re-employment wage is defined as the average monthly earnings an individual receives in the year after the UI spell has ended. Table 1 also reports some observable characteristics of unemployed job seekers. In the IZA ED data, individuals use roughly four to five search channels, where most individuals in the sample look for job advertisements, ask friends or relatives for jobs or use online search tools. Many individuals are also offered help from their local employment agencies. Table 1 shows that job seekers send out thirteen applications on average at the beginning of the UI spell. From the PHF dataset, we extract some information regarding assets (in particular, liquid assets) of the unemployed. In Table 1, we show different quantiles from the net liquid asset distribution of the unemployed in the sample. We see that asset holdings are indeed very heterogeneous, where nearly half of the individuals have hardly any assets.[3] By contrast, roughly 10% of individuals have more than 40,000 euros in liquid assets. Net assets, which also include real estate, are larger, on average. Finally, the JVS shows that firms receive, on average, fifteen applications and that it takes around two months to fill an open vacancy.[4]

1.2. *Descriptive Facts*

In this subsection, we provide some evidence on individuals' job-finding rates and search effort, as well as on firms' screening and interview decisions, which motivates our modelling assumptions. We then build a model of the job search and recruitment process, which incorporates these different aspects.

1.2.1. *Job-finding rates*

The job-finding rate of unemployed job seekers in Germany is shown in Figure 1(a). In the first months of unemployment, exit rates out of unemployment are above 10%. However, job-finding rates decrease throughout the spell and are only 5% after one year and 2.5% after two years of unemployment.[5] Hence, the chance to find a job diminishes the longer someone is unemployed. There are two explanations for this decline in the hazard rate from unemployment: first, selection/heterogeneity or, second, (true) duration dependence. Heterogeneity can enter in the form of the productivity differences of job seekers. Duration dependence describes the declining job prospects for individuals, given their type. Most likely, both selection and duration dependence contribute to falling hazard rates.

1.2.2. *Search effort*

Since we are interested in dynamic UI policies, it is important to consider how individuals change their search effort over the unemployment spell, because search effort responses are a main determinant of the moral hazard costs associated with UI. Figure 1(b) illustrates the number of applications that agents write per month as a function of their unemployment duration. At the beginning of the unemployment spell, they send, on average, more than thirteen applications per month, while after six months approximately nine applications are sent, and after twelve months only eight applications are sent. Hence, the average search effort seems to decrease over the unemployment spell.[6] Note that we have ignored other measures of search effort for now, for example the number of search channels or the time used in a job search. Our choice is motivated by the fact that our model explicitly allows agents to send out applications.[7]

1.2.3. *Multiple applications per vacancy*

An important factor that determines job search outcomes is how many other applicants are searching for a similar job. Hence, depending on the number of applications per vacancy, the job-finding rate might be higher or lower for a given search effort. The importance of these

[3] Net liquid assets are defined as the difference between liquid assets and short-term debt, such as credit card debt.

[4] This time is defined as the difference between the release of the job advertisement and the acceptance of a job offer by an applicant.

[5] The small spike at twelve months is due to the benefit exhaustion, which leads more individuals to exit unemployment. See DellaVigna *et al.* (2017) for a detailed exploration of the benefit exhaustion spike.

[6] Declining search efforts over the UI spell was also documented in the United States by Krueger and Mueller (2011).

[7] Lichter (2016) also used the number of applications as a search measure and discussed this choice in more detail.

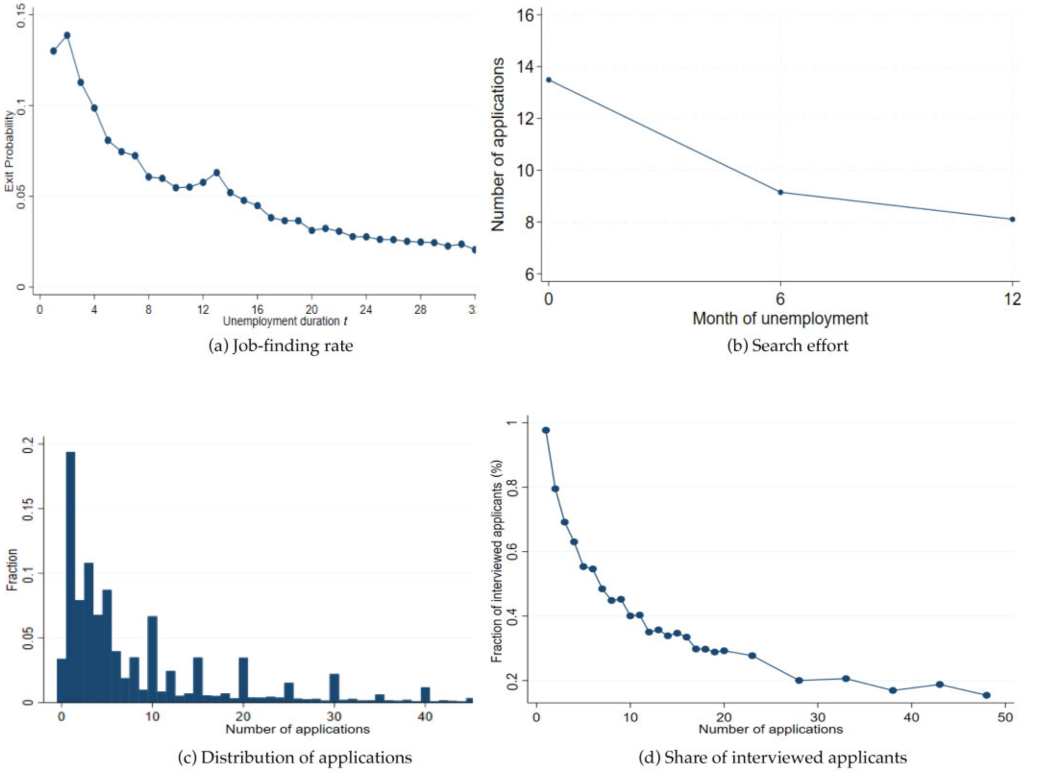

Fig. 1. *Descriptive Facts.*
Notes: Panel (a) shows the job-finding probability (hazard rate) of individuals on the *y* axis as a function of the unemployment duration on the *x* axis. *Source:* SIAB. Panel (b) shows the mean number of applications unemployed agents send out in the first month of unemployment, the sixth month of unemployment and after one year of unemployment. *Source:* IZA ED. Panel (c) illustrates the distribution of applications across vacancies. The *y* axis denotes the fraction of vacancies that receive a certain number of applications. *Source:* JVS. Panel (d) shows the fraction of interviewed applicants as a function of the number of applications received. *Source:* JVS.

crowding-out effects depend on the number of competitors in the application for a job. Intuitively, if there are many applicants per vacancy, some job seekers will receive no job offer and will need to continue their search. Figure 1(c) plots the histogram of the number of applications an open vacancy receives. The average number of applications is approximately fifteen, with a median of five applications per vacancy. This panel suggests that firms have considerable flexibility to select the best applicant and that the outside option of a firm is to screen or hire alternative applicants.

1.2.4. *Employer screening*

In the recruitment process, employers typically select a subset of the applications they receive and proceed with further screening. In panel (d) of Figure 1, we show that the share of applicants receiving an interview invitation depends on the number of applications a vacancy receives. It is clear that the more applications there are, the less likely it is to be invited to an interview. The interview shares are around 50% for vacancies with five applications (i.e., at the median), and only 30% for vacancies with fifteen applications (i.e., at the mean). The figure highlights that the presence of multiple applications per vacancy is an essential feature for studying screening; firms who only receive one application for their vacancy almost always interview the applicant.

In the job vacancy survey, employers are also asked whether they consider unemployed applicants, depending on the unemployment duration of the applicant. Conditional on stating that they consider unemployed applicants, only 75% of firms consider applicants with more than a few months of unemployment duration (Online Appendix Figure A2) and only 60% of firms consider applicants with more than twelve months of unemployment duration. Hence, only 60% of firms that are, in principle, willing to consider unemployed applicants are willing to accept long-term unemployed applicants. Online Appendix Figure A2 illustrates this graphically.

This survey evidence complements the experimental findings of Kroft *et al.* (2013), who found that the callback rate (interview invitation) of an application that was sent out to open vacancies strongly depends on the unemployment duration presented in the CV of the applicant. In fact, the probability of receiving a callback from an employer declines by roughly 50% over the unemployment spell. Note that declining callback rates can, in principle, also be generated by models of human capital depreciation. However, Kroft *et al.* (2013) demonstrated that the decline in the callback rate is much weaker when the unemployment-to-vacancy ratio is high. This finding is hard to rationalise with human capital depreciation, since human capital would depreciate independently of labour market conditions. Employer screening, on the other hand, predicts that unemployment duration is less informative about productivity under adverse labour market conditions, since individuals with high productivity also stay unemployed for a longer period. This is in line with the evidence provided by Kroft *et al.* (2013).[8]

2. Model

Our model is designed to capture the empirical patterns presented in Section 1.2. The model builds on previous literature, which studies optimal unemployment insurance in a setting with risk aversion, endogenous effort and savings (see, e.g., Lentz, 2009). We extend this work by incorporating firms' hiring decisions to account for the empirical patterns (as in, e.g., Jarosch and Pilossoph, 2019). The key feature of our model is that workers are heterogeneous in productivity and firms have to select candidates from a pool of multiple applications. Since productivity is only observed by workers, firms base their hiring decisions on the expected productivity of each worker, taking unemployment duration and a noisy signal about worker quality into account.

2.1. *Workers*

Time is discrete and each period corresponds to a month. To keep our model in line with the conceptual framework of Kolsrud *et al.* (2018), who theoretically characterised time-varying optimal UI schedules, we assume that in each period, a new generation of unemployed workers is born and lives for T periods. Workers who have been unemployed for t periods get UI benefits

[8] In addition, note that they found that the callback rate declines strongly within the first six months of unemployment and is essentially flat afterwards. If the decline in callback rates is mostly driven by human capital depreciation, one would expect a more gradual decline that also affects the long-term unemployed, since it is unlikely that human capital declines very rapidly within six months and then ceases to decline.

that depend on t:

$$b_t = \begin{cases} b_1 & \text{if } t \leq D, \\ b_2 & \text{if } t > D. \end{cases}$$

Thus, workers can get an initial level b_1 for up to D months and a level of b_2 afterwards.[9] Workers differ in their productivity π_j (known to them) and each generation of workers contains a share α_j of type $j = 1, \ldots, J$. The distribution of workers' productivity types is a beta distribution over the interval [0, 1]. In addition, each type has an exogenous initial level of assets, denoted $k_{0,j}$.

Employed workers only decide on the optimal level of consumption and savings. Their budget constraint is

$$c_t = Rk_t + (1 - \tau)w - k_{t+1},$$

where k_t and k_{t+1} are the asset levels in each period. The corresponding value function and budget constraint for duration $t < T$ are

$$V_t^e(k) = \begin{cases} \max_{k_{t+1} \geq 0}\{u(c_t) + \beta V_{t+1}^e(k_{t+1})\} & \text{if } t < T, \\ u(Rk + (1 - \tau)w) & \text{if } t = T. \end{cases}$$

Workers are risk averse and discount the future at rate β and the interest rate is given by R. There are no separations and employment is an absorbing state.[10] In addition, all workers face a no-borrowing constraint ($k_{t+1} \geq 0$).[11]

Unemployed workers decide on both consumption and savings and their search intensity. Searching with intensity s has a cost $\psi_t(s)$, but leads to a match probability $p(s) = s$, which can be interpreted as sending an application to a firm.[12] Note that we allow the search cost function to depend on t. In the estimation, we allow for *search decay*, capturing the idea that sending out applications becomes harder over time.

Importantly, the probability of exiting unemployment (that is, the hazard rate) contains both the probability of meeting a firm and of receiving a job offer from the firm:

$$h_{j,t} = s_{j,t} \cdot g_j(t). \tag{1}$$

Here, $g_j(t)$ is the expected offer probability and is determined in equilibrium, as will be discussed in the next sections.[13] Jobs start in the next period. The survival rate in unemployment (that is,

[9] Note that, in practice, the amount of unemployment benefit paid is often tied to the pre-unemployment wage. Because our model abstracts from wage heterogeneity, the pre-unemployment wage is conceptually indistinguishable from the post-unemployment wage.

[10] Allowing for separations is in principle possible, but would complicate the model by generating an endogenous initial asset distribution. Hence, for simplicity, we assume that jobs last forever.

[11] The no-borrowing assumption is standard in the literature (see, e.g., Chetty, 2006) and creates an insurance motive for the government in the first place. Without borrowing constraints, individuals would simply take out a loan and there would be no need for the government to provide insurance to the unemployed.

[12] Our model can be extended to include multiple applications per worker by assuming that workers who search with intensity s stochastically send out S applications, where S is drawn from a Poisson distribution with mean s (as in Kaas, 2010). It can be shown that the hazard rate here would be $h_j(t) = 1 - \exp[-g_j(t)s_{j,t}]$. In this case, computation of the offer probabilities would be more complicated and require simulation, as some workers who are offered jobs reject them in order to accept other offers. As this would also introduce an additional coordination friction, we focus on the simpler model with one application per worker.

[13] Note that our use of the term *offer* probability is identical to the term *hiring* probability in Lehr (2017) and Kolsrud *et al.* (2018).

the probability of still being unemployed after t periods) is then defined as

$$S_{j,t} = \prod_{t'=0}^{t-1}(1 - h_{j,t'}).$$

Taken together, the value function for unemployed workers for $t < T$ is given by

$$V^u_{j,t}(k) = \max_{s,k_{t+1}\geq 0} \left\{u(c_t) - \psi(s) + \beta h_{j,t}(s)V^e_{t+1}(k_{t+1}) + \beta[1 - h_{j,t}(s)]V^u_{j,t+1}(k_{t+1})\right\}.$$

The budget constraint is $c_t = Rk_t + b_t - k_{t+1}$. Note that changes to the benefit system influence the value of unemployment relative to employment and therefore affect workers' search decisions. In the terminal period, the value function is

$$V^u_{j,T}(k) = u(Rk + b_T).$$

2.2. *Firms*

2.2.1. *Matches, information structure and production*

When workers are matched with a firm, a match-specific productivity $q \in \{0, 1\}$ is drawn, which means that the worker is either suitable or unsuitable for the vacancy. The probability of being suitable is given by π_j, so that more productive workers have a higher likelihood of being suitable for any given match.[14] Whether an applicant is suitable or not is not known to the firm, but can be revealed through a job interview.

Firms produce an output y when employing a suitable worker and zero otherwise. Thus, conditional on being suitable, workers produce the same output. This assumption simplifies the model.[15] Workers are matched to firms according to an urn-ball-matching technology, where each matched worker randomly arrives at a firm. From the point of view of the firm, the number of applications it receives follows a Poisson distribution with parameter $\mu = a/v$, where a is the mass of matched workers and v is the mass of vacancies. For each candidate, firms do not observe if they are suitable, but only their unemployment duration and a noisy signal about the type of worker. Whenever a worker is suitable, their signal is drawn from $N(1, \sigma)$, whereas unsuitable workers send a signal from $N(0, \sigma)$. This captures the idea that firms do not perfectly observe whether a worker is suitable before conducting interviews, but there is some noise in the process and parameter σ governs the extent of this noise. Firms can interview applicants and thereby perfectly reveal their productivity. The goal of the firm is therefore to identify the candidates that are most likely to be suitable and to verify their suitability through the interview.

2.2.2. *Wages*

To keep the analysis tractable, we follow Jarosch and Pilossoph (2019) in assuming an exogenous wage.[16] In our model, firms pay an exogenous wage w to the applicant they eventually hire. One

[14] This is similar to the setup of Fernandez-Blanco and Preugschat (2018), who also assumed that workers differ in their probability of being suitable for vacancies.

[15] An interesting extension would be to allow the output between low- and high-type workers to differ; for example, by assuming that the match-specific productivity draws result from $N(\mu^P_j, \sigma^P)$ distributions, with the mean μ^P_j being higher for high-type workers. This would introduce an additional trade-off into the model because it would be possible to increase the aggregate output by providing firms with as much information as possible about the workers. In our setup, the planner prefers to eliminate statistical discrimination because it introduces risk for workers (of having very long periods of unemployment) and because some workers are inefficiently long-term unemployed.

[16] Jarosch and Pilossoph (2019) motivated this by assuming that workers have zero bargaining power in the Nash bargaining problem, so that their wage rate is always equal to the exogenous unemployment benefit b (the outside

way to interpret this is in the spirit of wage-posting models, in the sense that firms have previously committed to a wage (without modelling the determination of this wage explicitly) and then make their recruiting decisions conditional on this wage. Of course, in a fully spelled-out wage-posting model, the wage rate would be endogenous and change with market conditions. However, our point is that the basic decision structure for firms—selecting the most promising applicant for a given wage—would be similar in such a more complex setting.

Since this is conceptually an important assumption, it deserves further discussion (see also Online Appendix B). In particular, a fixed wage rules out that firms can offer lower wages to the long-term unemployed, so that unemployment duration would be reflected in both the job-finding rate and the wage. However, as also pointed out by Jarosch and Pilossoph (2019), there is little empirical evidence that this is an important concern. If lower wages compensate for the lower expected productivity, callback rates could be flat or even increase with duration: instead, the experimental studies find a decline. Furthermore, the decline of re-employment wages over the unemployment spell is mild (Fernandez-Blanco and Preugschat, 2018), which also holds in our German data (see Online Appendix Figures A4 and A5 for both re-employment wages and reservation wages). As a result, the assumption of a fixed wage is a reasonable first pass when studying the impact of screening. The main difficulty in incorporating endogenous wages is technical: on the one hand, random search models with multiple applicants per vacancy (such as ours or that of Jarosch and Pilossoph, 2019) make it harder to have endogenous wages.[17] On the other hand, directed search models (such as that of Fernandez-Blanco and Preugschat, 2018) have endogenous wages; however, it is harder to incorporate other crucial features when studying optimal UI, such as endogenous search effort and savings.

To ensure that our optimal policy results are not primarily driven by our assumptions on wages, we have experimented with several extensions, which we describe in more detail in Online Appendix B. One reason why wages could affect the analysis is that allowing wages to differ across types could differentially change the search incentives of high- and low-type workers, thereby influencing dynamic selection and the extent of statistical discrimination. In the extension, we allow for exogenous wage heterogeneity to capture this effect. We find that the main qualitative conclusions from our analysis are similar in this case (see Online Appendix B), which means that our main results are not driven by the assumption of homogeneous wages. In addition, we show that exogenously allowing for a mild decline in re-employment wages (as observed in the data; see Online Appendix Figure A4) leads to very similar results.

2.2.3. *Interviewing and hiring decisions*

Firms make their interviewing and hiring decisions while facing multiple applicants. We assume that there is a very small but positive screening cost C of conducting an interview ($C \to 0$). Therefore, the optimal firm behaviour is to rank applicants by their expected productivity (i.e., their likelihood of being suitable) and sequentially interview applicants until one applicant turns out to be suitable. The other applicants are not hired. Since the firm always has to pay the wage, it will never hire an unsuitable worker. A novelty in this setting, relative to previous literature on ranking by unemployment duration (e.g., Blanchard and Diamond, 1994; Fernandez-Blanco and

option). Note that, with time-varying benefits and endogenous effort, it is more difficult to motivate fixed wages through bargaining: workers' outside options change over time and zero bargaining power would also result in optimal search effort being zero, as workers would be indifferent between work and unemployment. Instead, the fixed wage could be thought of as (an extreme form of) market rigidity.

[17] This is due to random search models typically relying on Nash bargaining, which is a solution for one-to-one bargaining. However, with multiple applicants per vacancy, one firm would simultaneously bargain with multiple workers.

Preugschat, 2018) is that firms do not form their ranking based only on unemployment duration, but instead combine the information from both unemployment duration and the noisy signals by computing the expected productivity. Firms know the composition of the pool of applications, that is, on average, how many applicants with duration t are suitable ($\mathrm{ST}_t = \sum_{k=1}^{K} \alpha_k S_{tk} s_{tk} \pi_k$) and how many are not suitable ($\mathrm{NST}_t = \sum_{k=1}^{K} \alpha_k S_{tk} s_{tk} (1-\pi_k)$). These numbers take into account the facts that the survival rates differ between low- and high-type workers (S_{tk}) and that search effort can also differ (s_{tk}). For example, if high-type workers search more than low-type workers at a given duration t, more applications will come from high-type workers and there will therefore be more suitable candidates among applicants.

To evaluate the expected productivity (Π) of an applicant with duration t and signal ϕ, firms then combine these likelihoods of drawing a suitable applicant with the information contained in the signal by using Bayes' rule:

$$\Pi(t, \phi) = \frac{f_1(\phi) \cdot \mathrm{ST}_t}{f_1(\phi) \cdot \mathrm{ST}_t + f_0(\phi) \cdot \mathrm{NST}_t}. \tag{2}$$

Here, f_1 is the density of the normal distribution with mean 1 and variance σ and f_0 similarly is the normal density with mean 0 (recall that suitable applicants send a signal from f_1 and non-suitable ones from f_0). The equation captures that whether an applicant has a high expected productivity (that is, is highly likely to be suitable) depends on two factors. First, he/she can send a high signal. Second, when a large fraction of job seekers with duration t are suitable, firms infer that the applicant is also likely to be suitable. In the limit case $\sigma \to 0$, the signal perfectly reveals suitability and there is no reason to take the unemployment duration into account. This will serve as the full-information benchmark in the analysis. Conversely, when $\sigma \to \infty$, the signal contains no information and firms only rank applicants based on duration. For intermediate cases with $\sigma \in (0, \infty)$, firms weigh the information contained in both components and their relative importance is endogenous. For example, when the benefit system keeps productive types in the pool longer, unemployment duration can become less informative about productivity and the ranking order then depends more strongly on the signal.

Having defined $\Pi(t, \phi)$, we can now turn to the main object of interest from the firm side, which is the offer probability $g_j(t)$ (see (1)). When an applicant (of type j and duration t) sends his/her application, it arrives at a firm with a random number of applicants. Thus, we need to compute the probability that no other suitable applicant is ranked higher than this applicant and hired instead. This requires some algebra relating to urn-ball matching and the Poisson distribution, which is relegated to Online Appendix A. There, we show that the offer probability $g_j(t)$, conditional on being type j with duration t, is

$$\begin{aligned} g_j(t) &= \pi_j \int_{\phi} \exp[-p(\phi, t) \cdot \mu] dN_{1,\sigma}(\phi), \\ p(t, \phi) &= \sum_{\tilde{t}=1}^{T} \sum_{\tilde{j}=1}^{J} \frac{a_{\tilde{j},\tilde{t}}}{a} \cdot \pi_{\tilde{j}} \cdot P[\Pi(\tilde{\phi}, \tilde{t}) \geq \Pi(\phi, t) \mid \tilde{j}, t, \tilde{t}, \phi]. \end{aligned} \tag{3}$$

On an intuitive level, $p(t, \phi)$ captures the probability that, conditional on being unemployed for t periods and sending signal ϕ, a randomly drawn 'other' applicant from the pool of all applications is both suitable and sends a signal high enough so that the firm interviews (and subsequently hires) this other applicant. In other words, $p(t, \phi)$ is the probability of *not* being hired when competing against an alternative candidate from the pool of all applicants. Then, the equation of the offer probability $g_j(t)$ accounts for three aspects. First, an applicant does, not only compete against one other applicant, but on average there are μ applicants per vacancy. Second, the signal that the applicant sends is stochastic, so that we need to integrate over ϕ. Third, the equation takes the probability that the applicant is suitable (μ_j) into account, as being hired requires being suitable in the first place.

The callback rate has a similar expression:

$$C_j(t) = g_j(t) + (1 - \pi_j) \int_\phi \exp[-p(\phi, t) \cdot \mu] dN_{0,\sigma}(\phi).$$

This is the model analogue to recent audit studies that measure the decline in the callback rate (e.g., Kroft *et al.*, 2013), and represents the probability of being contacted and screened by an employer. The callback probability is the sum of the offer probability (those cases where an applicant is contacted and subsequently hired) and the probability of being contacted, but found unsuitable. Note that there are two components that give rise to the negative effect of unemployment duration on the callback probability. First, for a given agent with a high duration, $p(\phi, t)$ tends to be high, which means that the firm is likely to first interview and potentially hire one other randomly drawn applicant. This depends on how informative unemployment duration is about types and on the composition of the pool of applications: if the short-term unemployed search a lot, it is more likely that a random other applicant has a short duration and is potentially considered first. Second, this effect is scaled by the mean number of applications per vacancy, which is given by μ. In the extreme case of no competition ($\mu = 0$), the offer rate is flat and equal to π_j. In the case of a large applications-per-vacancy ratio μ the competition for jobs is large and callback rates are lower.

2.2.4. *Vacancy posting and free entry*

The mass of vacancies is pinned down by a free-entry condition. As in Lise and Robin (2017), firms can pay $c(v)$ to advertise v vacancies. Vacancies last for one period. The value of an additional vacancy is the net output multiplied by the probability of receiving at least one suitable application:[18]

$$J^v = \frac{y - w}{1 - \beta}\left[1 - \exp\left(-\frac{\sum \pi_j a_j}{v}\right)\right].$$

In equilibrium, the marginal vacancy costs are equal to the expected value of an additional vacancy:[19]

$$c'(v) = J^v.$$

Conceptually, free entry implies that firms can exit when posting vacancies becomes less profitable. Hence, vacancies might negatively or positively react to changes in unemployment policies. In our framework, different benefit schemes can reduce firm profits by either reducing overall search effort or by reducing the number of applications of high-type workers relative to low-type workers, because each case makes it less likely that vacancies receive at least one suitable candidate. As a result, firms would reduce the number of vacancies being posted. Later, when we

[18] Note that we assume that vacancies survive forever and that after the vacancy is filled it stays filled forever. This is a helpful approximation, especially when T is large enough.

[19] Depending on the functional form of $c'(v)$, vacancy creation rents accrue to firms if vacancy costs are not constant. However, it is not obvious how to interpret these rents and we ignore them throughout the rest of the paper.

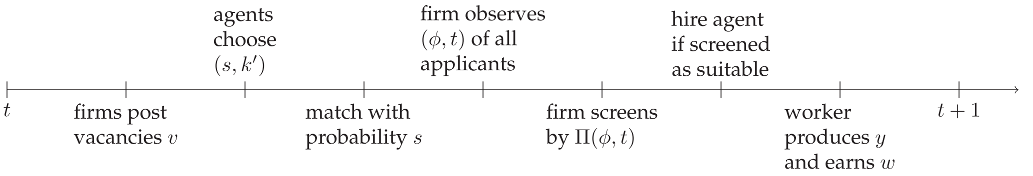

Fig. 2. *Timing of the Model.*

discuss optimal policy, these incentives for vacancies must be taken into account. In Figure 2, we summarise the timing of our model graphically.

2.3. *Equilibrium*

The equilibrium of the model consists of

- policy functions for search effort s_{j,t,k_t} and savings $k_{t+1} = g_u(k_t, t, j)$ for the unemployed and $k_{t+1} = g_e(k_t)$ for the employed, for each type j and duration t,
- survival functions $S_{j,t}$,
- expected offer rates $g_{j,t}$,
- a mass of vacancies v,

such that the policy functions of workers solve the problems described by the value functions for the employed and unemployed, and such that the expected offer rates are optimal according to (3), given the implied survival rates.[20]

3. Estimation

So far, we have described the data and some empirical facts, followed by a discussion of the model and the mechanisms. In this section, we connect our model to the data. We first present the estimation setup and then discuss the estimation results.

3.1. *Setup*

3.1.1. *Specification*

To estimate the model formulated in Section 2, we impose the following functional forms on the instantaneous utility function and the search cost function:

$$u(c) = \frac{c^{1-\gamma}}{1-\gamma}, \qquad \psi_t(s) = A_t \frac{s^{1+1/\lambda}}{1+1/\lambda}, \qquad A_t = \psi_1 t^{\psi_2}.$$

Here, λ denotes the elasticity of search effort with respect to the value of employment. The functional form is a common assumption and used in Lentz (2009) and DellaVigna *et al.* (2017). The instantaneous utility function is a standard CRRA utility function, where γ is the risk-aversion

[20] While the uniqueness of the equilibrium cannot be proved analytically, we checked for the possibility of multiple equilibria, especially around the estimated parameter values, and always converge to the same equilibrium.

parameter and, simultaneously, the inverse of the inter-temporal elasticity of substitution.[21] Parameters ψ_1 and ψ_2 capture the notion that the search efficiency can change over time, so that searching is harder for individuals with a longer unemployment duration.[22]

In our model, agents are heterogeneous in two dimensions: first, their probability of being suitable applicants and, second, their initial assets. In our baseline version of the model, we estimate a discrete approximation to the beta distribution with twenty types and assume that there are three different asset types, which in total leaves us with $J = 60$ types.[23] Signals are drawn from normal distributions with mean 1 when the worker is suitable and mean 0 when they are unsuitable.[24] We set initial assets for the unemployed to be uniformly distributed with 0, 500 and 3,000 euros. These values are set in order to roughly match the liquid assets of unemployed individuals in the PHF dataset. The wage received by agents during employment is fixed and we set $w = 1{,}605$ euros, which matches the mean re-employment wage in our sample of unemployed. The estimation is based on the current schedule, so that benefits b_t are set to a replacement rate of 63.5% within the first year and social assistance is equal to 40% after one year.[25] These numbers closely capture the benefits paid to the unemployed in our sample period. The time horizon in our model is $T = 96$, which amounts to eight years. By choosing this relatively large time horizon, we avoid agents' search behavior being influenced by end-of-life effects.[26] The vacancy posting costs are quadratic in the number of vacancies. The functional form for the vacancy posting costs we use is $c(v) = \kappa v^{1+\rho}$, where we set $\rho = 1$ to obtain quadratic vacancy costs. Note that, for the purpose of the estimation, the free-entry first-order condition (FOC) can be written as

$$\frac{1-\beta}{y-w}\kappa(1+\rho)v^{\rho} - 1 + \exp(-\mu_S) = 0.$$

We define an auxiliary variable as

$$c = \frac{1-\beta}{y-w}\kappa(1+\rho).$$

With this definition, the FOC becomes

$$cv^{\rho} - 1 + \exp[-\mu_S(v)] = 0.$$

Then, we estimate c (i.e., the ratio of κ and $y - w$) rather than each of its components individually, as it is only the ratio that matters for the vacancy posting.

[21] Alternatively, one could think about a CARA utility specification. The constant relative risk-aversion choice is motivated by the possibility of wealth effects, which implies different attitudes toward gambling with respect to wealth; that is, individuals who have less savings will search more. Shimer and Werning (2008) compared the implications of CARA and CRRA to optimal UI and found only minor differences, because wealth effects are quantitatively very small in a search model such as ours.

[22] In a simple manner, this captures various reasons for such 'search decay', such as psychological discouragement effects or fewer vacancies to apply for over time.

[23] Note that productivity and initial assets are uncorrelated.

[24] The means are a pure normalisation because we estimate the SD of the normal distribution.

[25] In Germany, after UI benefits run out, job seekers can apply for unemployment assistance (UA), which is means tested and a fixed payment that does not depend on the pre-unemployment wage. Hence, we choose a value for the replacement rate that roughly amounts to the replacement rate that a typical UA recipient would receive.

[26] Mechanically, in $T = 96$, agents stop searching because it only provides disutility to them. This *end-of-life* effect also influences search efforts in the previous periods. However, in our specification, these effects quickly become small and do not influence search behaviour in a quantitatively important manner in the first years of unemployment.

3.1.2. *Estimation*

Some additional parameters are set prior to estimation to standard values from the literature. We set the monthly time discount parameter equal to $\beta = 0.995$, which leaves us with an annual discount factor of roughly 5%. Risk aversion is equal to $\gamma = 2$ as in Chetty (2008) and Kolsrud *et al.* (2018). The interest rate is set to $R = 1/\beta$ as in Chetty (2008), Shimer and Werning (2008) and Lentz (2009). This leaves us with the following parameters to be estimated:

$$\theta = \{\lambda, B_1, B_2, c, \psi_1, \psi_2, \sigma\}.$$

Thus, the parameter vector contains the search effort elasticity λ, the two parameters of the beta distribution (B_1, B_2), vacancy posting parameter c, the search decay parameters (ψ_1, ψ_2) and the variance of the signal σ.

We estimate our model via minimum distance. This procedure minimises the distance between the data moments and their model equivalents. The theoretical moments can be computed analytically, given the solution of the model and the reduced-form moments are estimated as described in Section 1.2. We weight each group of moments (that is, the hazard moments, the search moments, the multiple spell moments and the applications-per-vacancy ratio) by the inverse of the number of moments in order to give comparable weights to each group.[27] In addition, the deviations are scaled in percentage terms for comparability. For the estimation of the parameters, we use a genetic algorithm, which is a global optimisation routine and can deal with potentially complex objective functions.

3.1.3. *Moments*

For the data moments, we combine information from each of the datasets that we discussed in Section 1. First, our moment vector includes the hazard moments from the first twenty-four months, grouped into six bins. We also include the change in the search effort in months 6 and 12 relative to the initial effort. Then we add the average number of acceptable applications that a vacancy receives, as seen in Figure 1.[28] Finally, we add six multiple spell moments where we use the mean unemployment duration in spell two, conditional on the unemployment duration in spell one. Note that we mimic the multiple spell sample in our model by simulating two unemployment spells for workers with the same type and the identical level of initial assets. This preserves the intuition of the length of the first unemployment spell being informative about the second spell of a certain type, while avoiding the explicit modelling of job destruction and keeping our framework more in line with standard UI frameworks.[29] Online Appendix Figure A3 shows this non-parametrically. This figure shows that the longer an individual's UI duration is in the first spell, the longer their UI duration in the second spell. As discussed in Alvarez *et al.* (2023), the stronger the correlation between the duration of unemployment in the two spells, the more important heterogeneity must be. The relatively small slope of the curve suggests that duration dependence might be important and that heterogeneity is not the sole driver of the declining hazard, as a high duration in the first spell implies a relatively small increase in the expected duration of the second spell. This leaves us with a total of fifteen moments to match.

[27] For example, with six hazard moments, each moment gets weighted by $\frac{1}{6}$. This ensures that the estimation does not give a disproportionate weight to fitting the hazard rates (as opposed to the applications-per-vacancy ratio for instance) simply because of the number of moments.

[28] To be precise, we truncate the moment at 250 applications; however, only a handful of firms report that number of acceptable applications.

[29] Empirically, we extend our sample to the period from 1983 until 2011, such that we have a sufficiently large sample of individuals with two unemployment spells.

Table 2. *Estimated Parameters.*

Parameter	Description	Estimated value
σ	Variance of the signal	8.955
λ	Search cost elasticity	0.461
c	Vacancy cost	0.813
ψ_1	Search decay, parameter 1	3.111
ψ_2	Search decay, parameter 2	0.481
B_1	Type distribution, parameter 1	6.144
B_2	Type distribution, parameter 2	6.635

Notes: This table summarises the estimation results of our parameters.

3.1.4. *Identification*

The parameters are jointly identified if any parameter vector θ has distinct predictions in terms of the model moments. Intuitively, changing a certain parameter needs to have different implications for the moment vector $m(\theta)$ compared with changing another parameter. While all parameters can influence all moments and are estimated together, we can discuss a few particular relationships between the parameters and the data moments. In our model, the level and slope of the hazard curve are strongly influenced by the parameters of the distribution of worker types, which determines how fast low- and high-type workers exit unemployment. The search cost elasticity λ is also closely related to the hazard curve. The decline in search effort also informs the estimation of the search decay parameters ψ_1 and ψ_2, as search decay means that the returns to search decline over the course of the unemployment spell. The multiple spell moments deliver additional information on the unobserved heterogeneity in the model and its relative importance vis-à-vis duration dependence. The higher the slope of the curve of the mean duration, the more heterogeneity in job-finding rates there should be. The intuition here is that the observation of two spells, in principle, allows us to estimate a fixed effect for individuals. If the correlation between the UI duration in spell one is strongly correlated with the UI duration in spell two, this suggests a large amount of heterogeneity (Alvarez *et al.*, 2023), and vice versa. This information is particularly helpful to estimate σ since the variance of the signal determines the importance of duration dependence in the model.

3.2. *Estimation Results*

In Table 2, we show the estimated parameters. We estimate the search cost elasticity λ to be 0.461. The parameters of distribution of worker types are best illustrated by plotting the implied beta distribution, which is shown in panel (b) of Figure 3. The distribution is centred around approximately 0.5 and most of the probability mass is in the middle of the distribution, with some very good and some bad types. Note that we discretise the distribution in practice using $K = 20$ points.[30] The heterogeneity in the productivity will translate into a heterogeneity in offer rates, as shown in panel (b) of Figure 4. The search decay parameters ψ_1 and ψ_2 are 3.111 and 0.481, determining the decrease in search efficiency over time. We estimate the variance of the signal to be equal to $\sigma = 8.955$, which implies that the productivity is relatively noisy. In other words, signals are comparatively uninformative and the ranking is strongly determined by unemployment duration. To comprehend the importance of the signal versus the importance of the unemployment duration, we compare the average offer rate in the estimated model to the

[30] We apply Kennan's method of discretising continuous distributions where each type has an equal probability mass.

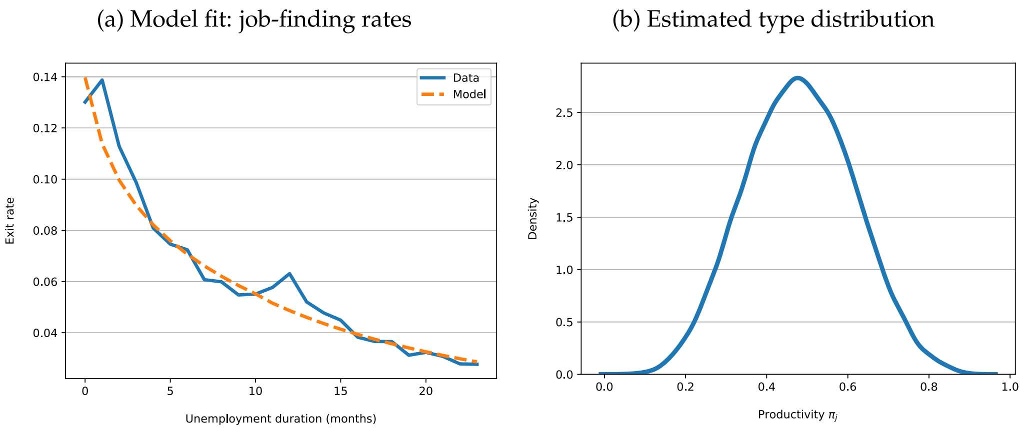

Fig. 3. *Estimation Results.*
Notes: Panel (a) shows the model-implied job-finding rates (dashed line) and compares them to the data (solid line). Panel (b) shows the estimated type distribution (which is a beta distribution). Note that in practice we discretise this distribution.

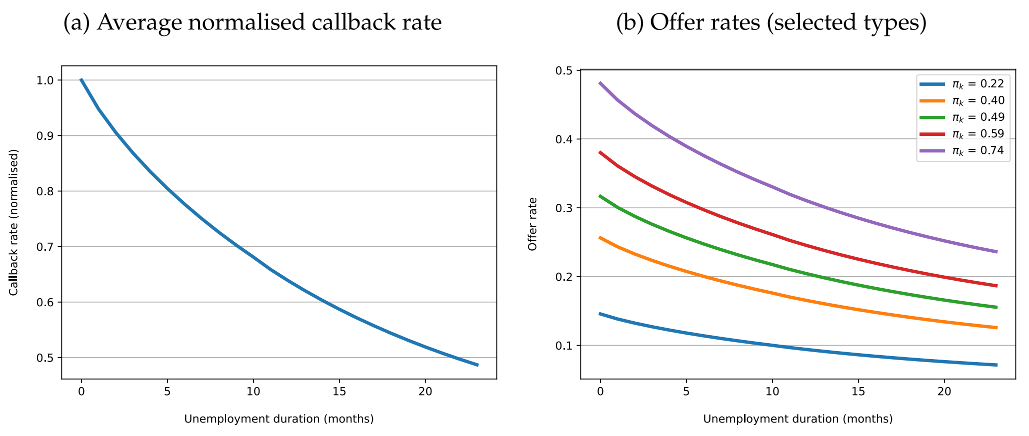

Fig. 4. *Model-Implied Callback and Offer Rates.*
Notes: Panel (a) shows the model-implied average callback rate of an application normalised to one in period $t = 1$. Panel (b) shows the type-specific offer rates for the unemployed that the model generates. The solid line corresponds to the low-type worker and the dashed line to the high-type worker.

hypothetical case where firms rank by duration only and disregard the signal ($\sigma \rightarrow \infty$). This is shown in Online Appendix Figure A6, which indicates that the presence of the signal affects offer probabilities even though the variance is relatively large.

In panels (a) and (b) of Figure 4, we illustrate the screening and hiring behaviour of firms that the model implies. Panel (a) shows the average decline in the callback rate of an application relative to period one. Our model suggests that the probability of being screened by a firm (that is, the probability of a callback) declines throughout the unemployment spell, and is only around 65% after one year of unemployment and drops below 50% after two years. The offer probability (panel (b)) is closely related to the callback probability, as it refers to the likelihood of both

Table 3. *Model Fit.*

	Data	Model
JFR (0–4)	0.120	0.111
JFR (5–8)	0.072	0.074
JFR (9–12)	0.057	0.057
JFR (13–16)	0.052	0.045
JFR (17–20)	0.036	0.037
JFR (21–24)	0.030	0.031
Decline in search effort, six months	0.678	0.694
Decline in search effort, twelve months	0.601	0.590
Mean duration, second spell (0–4)	0.118	0.123
Mean duration, second spell (5–8)	0.130	0.126
Mean duration, second spell (9–12)	0.139	0.128
Mean duration, second spell (13–16)	0.135	0.130
Mean duration, second spell (17–20)	0.138	0.132
Mean duration, second spell (21–24)	0.135	0.133
Applications-per-vacancy ratio	4.314	4.318

Notes: This table shows the fitted moments from our model. In the second column, one can see the data moments and in the third column the model-implied moments.

receiving a call and being suitable for the vacancy. For all types, offer rates decline because the screening probability declines. However, the offer probability per application of high-type workers is much greater than for low-type workers, reflecting the fact that high-type workers are more likely to be suitable. For all types, there is a substantial decline in the offer probability. For example, for the highest type shown in the figure, the offer probability declines by approximately half after two years of unemployment. The estimated heterogeneity and duration dependence in offer rates then translates to job-finding rates of agents. The job-finding rate is the product of the offer rate and the probability of sending out an application, namely, the search effort of the individual. The dashed line in panel (a) of Figure 3 shows the model-implied job-finding rate (JFR) of our model.

3.2.1. *Model fit*

How well does our model fit the targeted data moments and how well does our model describe non-targeted empirical patterns? In terms of targeted moments the fit is very good. Panel (a) of Figure 3 shows the fit of the hazard rate where the solid line is the data hazard and the dashed line is the model-implied hazard. The model fits the hazard curve closely, except for the spike around benefit exhaustion.[31] Table 3 shows the additional targeted data moments and the model-implied moments. Our model replicates the changes in search effort well, as well as the second spell moments, by capturing a positive slope. Finally, our model fits the applications-per-vacancy ratio very well.

These are two important pieces of evidence that we did not directly include in our estimation: first, callback rates and, second, duration elasticities with respect to potential benefit duration. Kroft *et al.* (2013) found in an experimental audit study that the callback rate from an application declines by about 40 percentage points after one year. In addition, the JVS data suggest that 40 percentage points of firms are not willing to consider unemployed applicants with an unemployment duration of one year or more, as shown in Online Appendix Figure A2. Our model indeed

[31] Here, other factors might be important; for example, individuals may exit registered unemployment because they are not eligible for social assistance. Because we do not model these features we disregard the spike at benefit exhaustion. See DellaVigna *et al.* (2017) for an exploration with present-biased and reference-dependent agents.

demonstrates a very similar pattern in terms of callback probabilities. As discussed above, our estimated model predicts a very similar average decline in callback rates. Thus, we are confident that the magnitude of the estimated screening channel in our model is plausible, since it compares well to the empirical findings on firm-induced duration dependence.

Schmieder *et al.* (2012) exploited quasi-experimental variations in age cutoffs of potential benefit duration in Germany. If an individual loses his/her job above a specific age cutoff, the maximal potential benefit duration increases from twelve to eighteen months. In their paper, the authors implemented a regression discontinuity design and found that an additional six months of benefits increases the mean non-employment duration by 0.78 months. In our model, we perform this simulation and we find that a benefit extension of six months implies an increase in the mean duration by 0.66 months. This is very close to the causal estimate from the data and increases our confidence that our estimated parameters imply a reasonable elasticity of search effort with respect to benefits. Furthermore, it ensures that the model-implied responsiveness to benefits is realistic. Since we are interested in optimal unemployment insurance, we want to have plausible behavioural patterns with respect to benefit payments.

4. Results

The goal of our analysis is to show how employer screening affects the optimal design of the UI system. We study this question by comparing the screening model to the 'full-information benchmark' (where firms observe whether an applicant is suitable and do not screen based on duration). The analysis is structured as follows. First, in Section 4.1, we compare labour market outcomes (that is, job-finding rates and their determinants) between the screening model and the full-information benchmark, since this is important for understanding why optimal policy differs between the two models. Second, in Section 4.2 we explain the optimal policy problem and discuss the optimal schedules. Third, Section 4.3 contains the main part of our analysis: we use a sufficient statistics framework to link the differences in labour market outcomes from Section 4.1 to the optimal policy results and explain the mechanisms behind why introducing screening changes the optimal schedule. Our key results in this regard are summarised in Section 4.3.2. Finally, Section 4.3.4 discusses in more detail how UI policy affects firms' hiring behaviour in the screening model and Section 4.4 discusses the welfare implications of taking screening into account when deciding about optimal policy.

4.1. *How Does Employer Screening Affect Labour Market Outcomes?*

We start by comparing labour market outcomes between the screening model and the full-information benchmark (FIB). In the FIB, we remove the information friction and assume that firms directly observe whether an applicant is suitable for a vacancy. This corresponds to setting the variance parameter of the signal (σ) to zero. As a result, firms hire a random suitable applicant. Since suitability is observed in this setting, there is no longer any reason for statistical discrimination based on unemployment duration. Figure 5 compares several important labour market outcomes between the screening model and the FIB (see also Table 4 for the underlying values and Online Appendix Figure A7 for heterogeneity by the productivity type).

Figure 5 shows that screening has a substantial effect on the job-search process. The job-finding rate after the first month of unemployment is almost doubled in the screening model compared to the full-information benchmark (14% versus 7.4%). Conversely, the job-finding rate

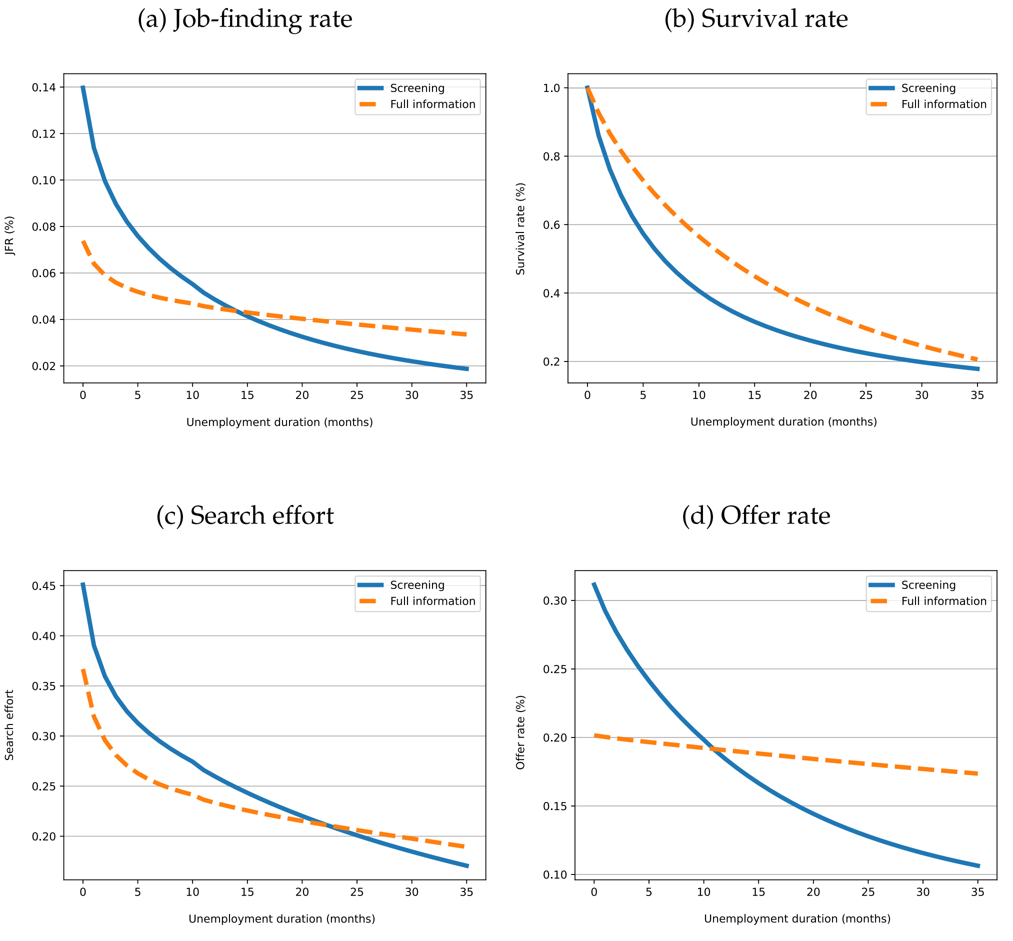

Fig. 5. *Job Search Outcomes under Screening versus the Full-Information Case.*
Notes: This figure shows average (a) job-finding rates, (b) survival rates, (c) search effort and (d) offer rates for the screening model and the full-information benchmark. Also see Table 4 for the underlying values.

Table 4. *Job Search Outcomes under Screening and the Full-Information Benchmark.*

		Offer rate	Survival	Search effort	Job-finding rate
One month	Full information	20.20	100.00	36.70	7.40
	Screening	31.10	100.00	45.10	14.00
Twelve months	Full information	19.10	53.90	23.60	4.60
	Screening	19.10	38.40	26.60	5.20
Twenty-four months	Full information	18.20	32.10	21.00	3.90
	Screening	13.40	23.70	20.80	2.90

Notes: This table shows average job-finding rates, offer rates and search effort at selected unemployment duration. These values correspond exactly to those shown in Figure 5. Note that all numbers are expressed in percentages. For the search effort, further recall that the match probability is normalised to $p(s) = s$, so that the value of effort in the table directly corresponds to the match probability (in percentages).

of the long-term unemployed is lower with screening (2.9% versus 3.9%). In our model, these effects are due to both changes in search effort and offer rates. Search effort in the beginning of the spell rises due to screening as workers try to escape the prospect of being less employable once they have been unemployed for a period of time, which contributes to the initial rise in the job-finding rate. At twenty-four months of unemployment, the search effort is slightly lower in the screening model (20.8 versus 21). Offer rates similarly increase in the beginning of the spell, as firms statistically discriminate and favour the short-term unemployed over the long-term unemployed (31.1% versus 20.2%). This also contributes to the rising job-finding rate at the start. The converse of this effect is that offer rates fall for the long-term unemployed; in the screening model, the offer rate after twenty-four months is 13.4%, whereas it is 18.2% in the FIB. Note that these values refer to the *unconditional* offer rate (averaging across all types). Therefore, offer rates also fall with unemployment duration in the full-information benchmark, which reflects the selection of job seekers by duration. As the composition of the unemployment pool deteriorates over time, less employable individuals remain, who also have a lower offer probability in the first place. By contrast, the *conditional* offer rate for each type is constant with duration in the FIB, while it falls in the screening model (see Online Appendix Figure A7). Because of the effects on both offer rates and search effort, duration dependence is much more pronounced with screening. For example, in the screening model, the average job-finding rate after two years is only 21% of what it was initially (2.9% versus 14%), whereas it falls to 53% of its initial value in the full-information benchmark (3.9% versus 7.4%).

It is worth comparing our results to those of Jarosch and Pilossoph (2019). As experimental studies only measure the decline in the *callback* rate, but not whether an applicant is ultimately hired, it is ex ante unclear to what extent declining callback rates also map into declining offer probabilities. Given that it is fundamentally difficult to measure offer rates in experiments (since this would require sending study participants to job interviews), quantifying the link between callback and offer probabilities requires a structural model of the hiring process, which links the available data to hiring outcomes. Jarosch and Pilossoph (2019) found that a large fraction of those who are not invited for interviews are low-type workers who would not have been hired anyway. Their results suggest that screening has a relatively small impact on actual job-finding rates; in their model job-finding rates are by at most 8% higher in the full-information benchmark than in the screening model (see Table 5 in their paper). Comparing this to Table 4, we find a larger impact on job-finding rates. Our results therefore point to a larger role of employer screening for aggregate labour market outcomes.

4.2. *The Policy Problem*

To determine the optimal unemployment policy, the government maximises welfare while balancing a lifetime budget constraint. We focus on two-step schedules that, as previously introduced in the model section, consist of three policy variables (b_1, b_2, D).[32] The proportional income tax τ is collected from the employed to finance the expenditures. In this case, the tax also has the interpretation of an actuarial fair insurance premium.

[32] While it is possible to imagine more complex specifications, such as having more steps or a fully unrestricted schedule, this setup has two key advantages. First, it closely corresponds to the UI policies that are in place in many countries (including, for example, the United States, Germany and Sweden). Therefore, the results can easily be related to the actual policies observed in these countries. Second, having only three policy variables makes the optimisation problem more tractable and allows us to reliably find a global maximum, even in the presence of a potentially complex objective function.

Table 5. *Optimal Schedule: Screening versus the Full-Information Benchmark.*

	Model	
Policy variable	Full-information benchmark	Screening
Initial benefit level (b_1)	59%	63%
Final benefit level (b_2)	46%	46%
Potential benefit duration (D)	19	26
Tax (τ)	22.8%	21.3%

Notes: This table shows the optimal schedule in the screening model and in the full-information benchmark.

More formally, the objective of the planner is to maximise the value of a newly born generation of unemployed.[33] We assume that every unemployed individual has the same welfare weight when born, which amounts to a standard utilitarian welfare criterion, as in Chetty (2006):

$$W(b_1, b_2, D, \tau) = \sum_{k=1}^{K} \alpha_k V_j^u(b_1, b_2, D, \tau).$$

However, the government can only maximise the welfare of agents subject to the following life-time budget constraint:

$$G(b_1, b_2, D, \tau) = \underbrace{\sum_{t=1}^{T} R^{-t}(1 - S_t) w\tau}_{\text{expected revenue}} - \underbrace{\sum_{t=1}^{T} R^{-t} S_t b_t}_{\text{exp. expenditure}}.$$

The government collects $w\tau$ and spends b_t on the unemployed and the probability of still being unemployed after t periods is given by the survival rate S_t. Thus, taken together, the problem of the government is

$$\max_{b_1, b_2, D, \tau} W(b_1, b_2, D, \tau) \quad \text{such that} \quad G(b_1, b_2, D, \tau) = 0.$$

To compute the optimal schedule, we numerically evaluate the social welfare function $W(P)$ on a grid for the benefit levels and potential benefit duration ($\{b_{1i}, b_{2i}, D_i\}_{i=1,\ldots,N}$), and select the schedule that yields the highest level of welfare.[34]

Table 5 shows the optimal schedules for both the full-information benchmark and the screening model. The main results are that in the screening model, the initial level of benefits (b_1) is higher (63% versus 59%) and the potential benefit duration (D) increases by seven months (nineteen versus twenty-six months). Despite the higher generosity, the UI system is less costly in the screening model and requires a tax of only 21.3% as opposed to 22.8% in the FIB. This is consistent with the observation that screening induces faster exit rates from unemployment due to

[33] Following previous work on employer screening and optimal UI (Lehr, 2017; Kolsrud *et al.*, 2018), we abstract from firms' profits in the social welfare function. A simple way of motivating this assumption is that firm profits disproportionately accrue to individuals at the top of the income distribution, who carry a lower social welfare weight or have a much lower marginal utility from consumption than the unemployed.

[34] For the two benefit levels, we set up the grid using 1-percentage-point steps. To ease the computational burden of solving the model for each grid point, we parallelise these function evaluations on a high-performance computing cluster. The advantage of this grid search approach is that it guarantees finding a global maximum without making assumptions on the shape of the objective function. In Online Appendix E, we show that the global maximum is robust (in the sense that the first and second best policies are very similar, for example), and that the objective function is smooth and concave for univariate variations of the parameters.

the additional incentive to search early in the unemployment spell (Figure 5). As we will explain in more detail in the next section, the changes in the optimal policy should be interpreted from the perspective of Figure 5: introducing screening affects firms' hiring behaviour and workers' search effort (relative to the full-information case), which changes aggregate job-finding and survival rates and therefore the optimal policy.

To conclude this section, it is interesting to relate these results to Kolsrud *et al.* (2018). They found empirically that introducing an inclining tilt ($b_2 > b_1$) into the schedule would be welfare improving in the context of a partial equilibrium model without employer screening.[35] To what extent does screening reinforce the case for an inclining schedule? We have also computed the optimal schedule in the screening model when holding the potential benefit duration fixed at $D = 19$. In this case the optimal schedule is $(b_1, b_2) = (0.64, 0.48)$. This means that the tilt becomes more declining in the screening model, since the planner mainly raises b_1. From the perspective of the sufficient statistics framework that we discuss in Section 4.3 below, one of the reasons for this result is that the moral hazard costs are higher for raising b_2 than for raising b_1. Therefore, the hiring externality does not reduce the moral hazard costs as much (in percentage terms) as it does for b_1.[36]

4.3. *Why Does Screening Change the Optimal Policy?*

The goal of this section is to analyse in more detail why the optimal schedule from the full-information benchmark is not optimal anymore in the screening model. As a starting point for this analysis, we compare worker and firm side outcomes between the screening model (*1*) after imposing the optimal schedule from the full-information benchmark on it and (*2*) when using the schedule that is actually optimal in the screening model. This comparison is shown in Table 6. We impose the optimal schedule from the FIB onto the screening model by setting $(b_1, b_2, D) = (0.59, 0.46, 19)$ and letting the tax rate adjust to balance the budget, given the job-finding and survival rates of the screening model.

On the worker side, the table shows that the optimal schedule first raises the time spent in unemployment, as the survival rates after both twelve and twenty-four months rise. For example, with the optimal schedule, 44.1% of workers are still unemployed after twelve months, whereas it is only 41.4% under the schedule from the FIB. In line with the increase in survival rates, aggregate search effort declines, reflecting the fact that more generous benefits discourage search effort. To illustrate how the selection of types changes with the optimal schedule, the table also shows a breakdown of survival rates by types (using as examples the lowest- and highest-type workers in the population). The lowest-type worker has a very low probability of being suitable for vacancies and is therefore likely to become long-term unemployed, whereas few of the highest-type workers enter long-term unemployment. With the optimal schedule, survival rates of both types increase, although more strongly for the high-type workers.[37] As we will

[35] Extrapolating from their sufficient statistics estimates using a structural model, they found that welfare would be maximised at $b_1 = 0.48$ for the short-term unemployed and $b_2 = 0.68$ for the long-term unemployed.

[36] For the case of b_1, we find that the externality-adjusted moral hazard costs are 49.7% lower in the screening model than in the full-information benchmark, which increases the optimal value for b_1. For b_2, the adjusted MH costs are only 24.6% lower in the screening model, so that the optimal level of b_2 does not increase as much. For screening to push towards a flat or inclining benefit profile, the externality from raising b_2 ($\text{EXT}(b_2)$) would need to grow relative to $\text{EXT}(b_1)$. Kolsrud *et al.* (2018) found that shifting the tilt towards a more inclining profile is welfare improving when $\text{CS}_{b_1}/\text{MH}_{b_1} < \text{CS}_{b_2}/\text{MH}_{b_2}$.

[37] The ratio of high-type workers relative to low-type workers at $t = 12$ is 0.36 with the schedule from the FIB and 0.4 with the schedule from the screening model.

Table 6. *Comparison of Schedules in the Screening Model.*

	Optimal schedule from the FIB	Optimal schedule from the screening model
Avg survival (twelve months)	41.4%	44.1%
Avg survival (twenty-four months)	26.9%	28.7%
Avg survival (twelve months), low type	68.4%	70.0%
Avg survival (twenty-four months), low type	57.0%	58.5%
Avg survival (twelve months), high type	24.7%	27.8%
Avg survival (twenty-four months), high type	11.2%	12.8%
Aggregate search effort	4.317	4.262
Vacancies	1.04	1.037
Applications per vacancy	4.152	4.11
Avg offer rate (one month)	31.4%	31.2%
Avg offer rate (twelve months)	19.8%	20.5%
Avg offer rate (twenty-four months)	14.6%	15.5%
Tax	18.8%	21.3%

Notes: In this table, we compare different outcomes of the screening model for (*1*) the optimal schedule from the full-information benchmark (i.e., setting $(b_1, b_2, D) = (0.59, 0.46, 19)$) and (*2*) the schedule that is optimal in the screening model ($(b_1, b_2, D) = (0.63, 0.46, 26)$). For the breakdown of survival rates by type (rows 3–6), we show the two examples of the lowest type and the highest type.

analyse in Section 4.3.4 below, these shifts in the composition of the unemployment pool will improve firms' beliefs about the long-term unemployed. Finally, regarding the tax, implementing the optimal UI schedule requires raising the tax relative to using the schedule that is optimal in the full-information benchmark. If the optimal schedule from the FIB was implemented in the screening model, the budget-balancing tax rate would only be 18.8%. The optimal schedule from the screening model then increases the UI budget, so that the tax is 21.3%. Note that the optimal schedule from the FIB is cheaper in the screening model than in the FIB (18.8% versus 22.8%). The reason for this is that workers exit unemployment faster, so that the government needs to collect less tax revenue.

On the firm side, the number of vacancies declines only slightly, so that the mean number of applications per vacancy overall declines (from 4.152 to 4.11). Regarding firms' profits, the assumptions from the estimation allow us to compute the change—rather than the levels—of profits (see Online Appendix A3 for details) between the policy regimes. We find that profits decrease very slightly by 0.53 percentage points. Importantly, switching to the optimal schedule increases the average offer probabilities at twelve and twenty-four months by around 1 percentage point (6.1% relative to the baseline value of 0.146), which means that it becomes more likely that long-term unemployed job seekers exit unemployment. We examine this increase and the mechanisms behind it more closely in Section 4.3.4 below.

4.3.1. *Sufficient statistics: background*

The social planner problem that determines the optimal UI schedule is complex, as it depends on various factors, such as the marginal utilities in unemployment and employment, how likely it is to be unemployed at different periods, the responsiveness of the search effort to benefits (that is, the corresponding elasticities) and firms' offer probabilities. Therefore, it is not easy to understand why a different schedule is optimal in the screening model compared to the full-information benchmark, as all of the factors are (potentially) different in the screening model. To better understand these differences, we connect our analysis to the 'sufficient statistics' approach from the public economics literature, which has developed simple formulae to capture the economic motives for and against providing UI benefits. The sufficient statistics help to show more clearly why screening changes the optimal policy and give some direct insights into the trade-offs of the social planner.

For this exercise, we first need to discuss the theory behind the optimal policy problem. We exactly follow the framework of Kolsrud *et al.* (2018), and a full technical derivation of the sufficient statistics can be found in their paper. Kolsrud *et al.* (2018) showed that the optimal schedule balances the *consumption smoothing gain* and the *moral hazard cost*. In addition, in employer screening models, there is also a hiring externality that reflects the fact that changes in the UI policy can change firms' interviewing and hiring decisions that arise in equilibrium. In the following, we briefly introduce the formulae and explain the intuition that is important for our analysis (for more details, see also Online Appendix A3). Since the optimal schedule in the screening model has both a higher initial benefit level b_1 and a higher potential benefit duration D (see Table 5), we need two sets of sufficient statistics—one for each of these policy variables.

***Sufficient statistics for b_1*:** We start by discussing the case of b_1 (the formulae for D will be very similar). We consider a simplified policy problem where the government only varies b_1 to find its optimal level (conditional on D and b_2):

$$\max_{b_1,\tau} W(b_1, \tau) \quad \text{such that} \quad G(b_1, \tau) = 0. \tag{4}$$

This problem is effectively univariate, as, conditional on b_1, the tax needs to be set to balance the budget. The first-order condition for b_1 can be written in terms of the consumption smoothing (CS) benefit, the moral hazard (MH) cost and the externality (EXT) term:

$$\text{CS}(b_1) = \text{MH}(b_1) - \text{EXT}(b_1). \tag{5}$$

Kolsrud *et al.* (2018) defined the *externality-adjusted* moral hazard cost as $\widetilde{\text{MH}} = \text{MH} - \text{EXT}$.[38] Equation (5) corresponds exactly to the first-order condition of problem (4) and means that at the optimal level of b_1, the insurance gain (CS) is equal to the externality-adjusted moral hazard cost. Therefore, when the goal is to understand why the planner picks a certain schedule, it is helpful to look at the sufficient statistics, as they give some direct insights into the trade-off the government faces.

Two of the key variables of the social planner problem measure the time individuals spend on average in the first and the second steps of the schedule. By $D_1 = \sum_{t=1}^{D} R^{-t} S_t$ we denote the (discounting-adjusted) time spent receiving b_1 and, similarly, $D_2 = \sum_{t=D+1}^{T} R^{-t} S_t$ refers to the time spent receiving b_2: the higher survival probabilities (S_t) are, the higher D_1 and D_2. The consumption smoothing gain is defined as

$$\text{CS}(b_1) = \frac{\sum_{t=1}^{D} \beta^t S_t E[u'(c_t^u)]/D_1 - \lambda}{\lambda}.$$

The interpretation is that CS captures the mean expected marginal utility in unemployment, weighted by the relative time unemployed individuals spend at each duration ($\beta^t S_t / D_1$), and compares it to the marginal value of public funds (the Lagrange multiplier λ on the budget

[38] Note that, in principle, the externality could also be added to the consumption smoothing gain (that is, on the left-hand side of the equation). We follow the discussion in Kolsrud *et al.* (2018) by adjusting the MH costs rather than the CS gain.

constraint from problem (4)), which can be thought of as the marginal utility of employed workers.[39] Therefore, the CS gain is commonly referred to as the gap in marginal utilities.

The moral hazard cost is

$$\mathrm{MH}(b_1) = \frac{\epsilon_{D_1,b_1} D_1(\tau + b_1)/b_1 + \epsilon_{D_2,b_1} D_2(\tau + b_2)/b_1}{D_1}. \tag{6}$$

The moral hazard term measures the importance of behavioural responses (represented by the elasticities ϵ_{D_1,b_1} and ϵ_{D_2,b_1}), which change the time individuals spend in unemployment, relative to the mechanical revenue loss from increasing b_1. The mechanical effect is the loss in revenue from having to pay higher benefits if behaviour and therefore survival rates are constant. To understand the expression for the MH cost, it is helpful to note that the impact of a raise in b_1 on the budget can be written as $\partial G(P)/\partial b_1 = -\mathrm{MRL} - \mathrm{BR}$, where MRL is the mechanical revenue loss and BR the behavioural response. The moral hazard term is then defined as $\mathrm{MH} = \mathrm{BR}/\mathrm{MRL}$. The MRL is equal to D_1 for the case of raising b_1 and the BR is the numerator from (6). The interpretation is that raising b_1 by 1€ reduces revenue by $(1 + \mathrm{MH})\mathrm{MRL}$€. The MH cost therefore indicates how strong the behavioural effects are *on top* of the mechanical effect from raising b_1.

Finally, the externality term depends on how firms adjust their interviewing and hiring decisions in response to changes in b_1:

$$\mathrm{EXT}(b_1) = \sum_{t=1}^{T}\sum_{k=1}^{K} \beta^t \alpha_k S_{t,k} s_{t,k} \frac{\partial g_{t,k}}{\partial b_1}(V^e_{t,k} - V^u_{t,k})\frac{1}{D_1\lambda}.$$

The externality term is new in screening models and would not be present in partial equilibrium models without firms. As a result, it is worth discussing this term in more detail. The externality term captures the notion that firms change their interviewing and hiring decisions after a UI policy reform. For example, when a UI policy changes the selection of types in a way that more productive types are among the long-term unemployed, this will increase offer rates for individuals with a high unemployment duration. As firms do not choose offer rates to maximise worker utility, these changes in $g_{t,k}$ have a first-order impact on welfare rather than dropping out due to the envelope theorem. Looking at the term more closely, the externality measures to what extent increases in b_1 can raise the job-finding rate via offer rates ($s_{tk}\partial g_{t,k}/\partial B$), which can increase welfare by moving workers from unemployment into employment ($V^e_{t,k} - V^u_{t,k}$).[40] These changes are weighted by the likelihood of each type staying unemployed for long enough to experience the increase in offer rates ($\alpha_k S_{t,k}$).

In practice, computing the sufficient statistics requires a value for the Lagrange multiplier. Since the planner problem in (4) is a one-dimensional optimisation problem, we can numerically solve for the optimal level of B and then compute the implied value of the Lagrange multiplier from the first-order condition at the optimum (see (5)).

***Sufficient statistics for D*:** Finally, we need sufficient statistics for the potential benefit duration (D), as our goal is also to understand why the planner raises D in the screening model relative to the full-information benchmark. The formulae are very similar to the case of b_1. As D is a

[39] Technically, the Lagrange multiplier can be obtained from the FOC for the tax rate and is a function of marginal utility of employed workers. See Online Appendix F for more details.

[40] Note that $s_{tk}\partial g_{t,k}/\partial B$ is the derivative of the job-finding rate with respect to B when the search effort is constant. Dividing by λ at the end simply results from rearranging the FOC, so that the externality term is on the right-hand side together with MH.

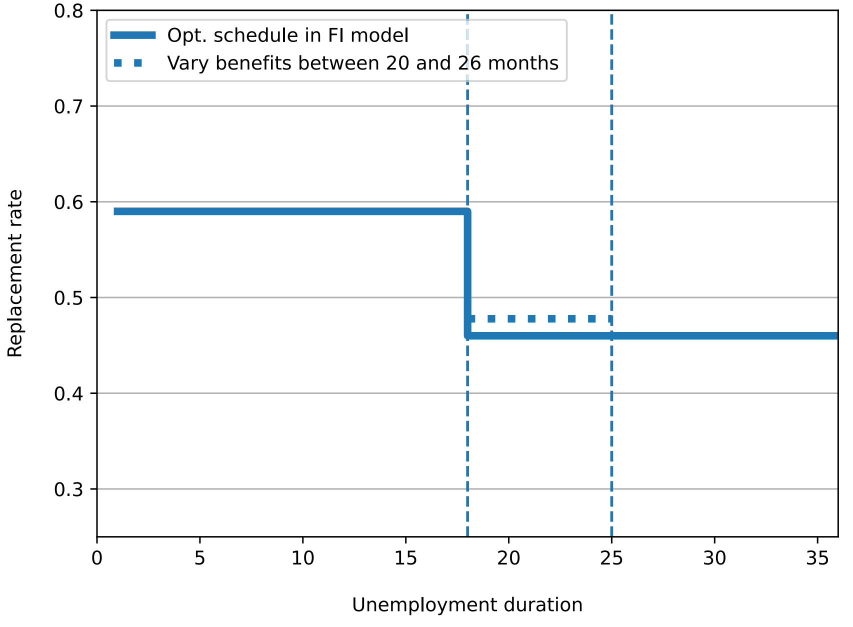

Fig. 6. *Sufficient Statistics for the Potential Benefit Duration (D).*
Notes: The figure illustrates how we derive sufficient statistics that capture an increase in the potential benefit duration (D).

discrete variable, we cannot take the derivative with respect to D directly. Instead, we consider the case of changing benefits between durations of 20 and 26 months (knowing that $D = 26$ is the optimal solution in the screening model). When setting benefits on this interval (which we denote b_D) from b_2 to b_1, this corresponds exactly to extending D to 26. For other values, it effectively creates a three-step schedule (see Figure 6). The purpose of this exercise to study why the planner finds it optimal to increase benefits on this interval due to screening.

We define $d_0 = 19$ and $d_1 = 26$. Given the three-step schedule, time spent in unemployment is now broken down into the three variables $D_1 = \sum_{t=1}^{d_0-1} R^{-t} S_t$, $D_2 = \sum_{t=d_0}^{d_1} R^{-t} S_t$ and $D_3 = \sum_{t=d_1+1}^{T} R^{-t} S_t$. The new formulae for the CS gain and MH cost are (EXT is unchanged)

$$\mathrm{CS}(b_D) = \frac{\sum_{t=d_0}^{d_1} \beta^t S_t E[u'(c_t^u)]/D_2 - \lambda}{\lambda},$$

$$\mathrm{MH}(b_D) = \frac{\epsilon_{D_1,b_D} D_1(\tau + b_1)/b_D + \epsilon_{D_2,b_D} D_2(\tau + b_D)/b_D + \epsilon_{D_3,b_D} D_3(\tau + b_2)/b_D}{D_2}.$$

4.3.2. *Sufficient statistics: overview of main results*

With these theoretical concepts available, we now discuss the results. The advantage of the sufficient statistics is that they reduce the complexity of the social planner problem to a few interpretable statistics. How does screening affect the insurance gain and the efficiency cost of providing UI benefits? Furthermore, how important is the hiring externality relative to these factors? We address these questions in this section.

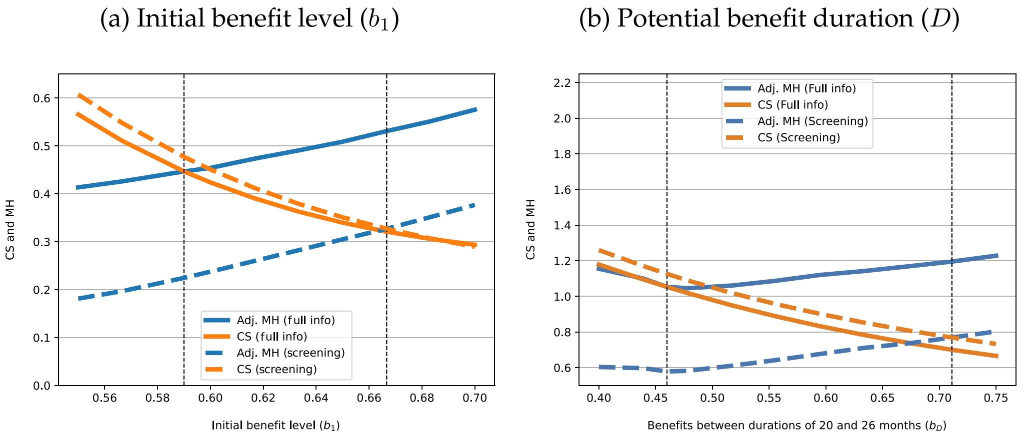

Fig. 7. *Sufficient Statistics: Why Does Screening Affect the Optimal Schedule?*
Notes: This figure shows the values of the CS gain, the MH costs, the hiring externality (EXT) and the externality-adjusted MH costs for different values of b_1 (panel (a)) and b_D (benefits between durations of 20 and 26 months, shown in panel (b)). The solid lines show the sufficient statistics in the full-information model and the dashed lines show them in the screening model, illustrating why the introduction of screening induces the planner to choose a different optimal policy. The vertical lines show the intersections of CS and $\widetilde{\text{MH}}$ (i.e., the optimal level).

Table 7. *Sufficient Statistics for b_1 Evaluated at $b_1 = 0.59$.*

	Full-information benchmark	Screening model	Difference (relative to the FIB)
CS	0.450	0.480	0.03 (6.8%)
MH	0.470	0.400	−0.07 (−14.1%)
EXT	0.020	0.180	0.16
$\widetilde{\text{MH}}$ (= MH − EXT)	0.450	0.220	−0.22 (−49.7%)

Notes: This table shows the sufficient statistics for the case of raising b_1 evaluated at $b_1 = 0.59$. The values correspond to those shown in Figure 7. It is optimal to increase b_1 whenever $\text{CS}(b_1) > \widetilde{\text{MH}}(b_1)$, and vice versa.

Figure 7 compares the sufficient statistics between the screening model and the full-information benchmark. In both models, we set $(b_2, D) = (0.46, 19)$ (the optimal values from the full-information benchmark) and vary the value of b_1. In each case, we let the tax adjust to balance the budget. Panel (a) shows the case of raising b_1. In the full-information benchmark, $b_1 = 0.59$ is optimal and therefore the consumption smoothing gain and the externality-adjusted moral hazard costs (the solid line) intersect at this point. Next, the figure shows that introducing screening shifts the consumption smoothing gain upwards and the adjusted moral hazard costs downwards (the dashed lines). The new intersection occurs at b_1 just above 0.66. Therefore, it is optimal to raise b_1 in the screening model.[41] Table 7 shows the values of the sufficient statistics at $b_1 = 0.59$. In the screening model, the consumption smoothing gain is 6.8% higher and the moral hazard cost is 14.1% lower. The externality plays a significant role, as it amounts to roughly 50% of the moral hazard cost. Panel (b) of Figure 7 shows that the reasons for raising D are similar: again, the consumption smoothing gain rises, while the moral hazard costs fall.

[41] Note that this calculation assumes that the only policy variable of the planner is b_1, and D and b_2 stay at their initial values. As a result, the optimal value of b_1 is higher than in Table 5, where the planner simultaneously increases D.

The main part of our analysis is to use these sufficient statistics results to analyse the mechanisms explaining why introducing screening changes the optimal policy. We begin by describing a high-level overview of the chain of events that makes it optimal for the planner to increase b_1. This is a summary of the main mechanisms, which will then be explained in more detail in Section 4.3.3 below. As the reasons for why the planner raises D are very similar to raising b_1, we focus on the case of raising b_1 in the main text. Online Appendix F contains a similar analysis for why the planner raises D. The chain of events once screening is allowed for (as opposed to being in the full-information benchmark) is as follows.

(1) Employers start screening by unemployment duration. Since firms discriminate against individuals with a long unemployment duration, offer rates are higher for the short-term unemployed and lower for the long-term unemployed, compared to the full-information benchmark (see Figure 5(d)).

(2) Workers internalise these changes and adjust their search effort. They search more at the beginning of the spell to avoid long-term unemployment.

Because of points (1) and (2), workers exit unemployment faster than they would in the full-information case. The average time workers spend receiving b_1 ($D_1 = \sum_{t=1}^{D} R^{-t} S_t$) falls significantly. Duration D_2 also falls. See panels (a)–(c) of Figure 5.

(3) These changes in offer rates and search effort already have implications for the planner in that they raise the MH cost and CS gain (note that there will be an offsetting effect on the MH cost; see point (4)(b) below).
   (*a*) Since workers exit faster and spend less time receiving b_1 (i.e., D_1 falls), it is cheaper for the government to raise benefits: both the MRL and the BR fall. The moral hazard cost is defined as the ratio between the BR and MRL. Since the MRL falls more strongly, this effect increases the MH cost.
   (*b*) In addition, because workers internalise screening, their search effort becomes more responsive with respect to benefits (the partial equilibrium elasticities $\epsilon^{PE}_{D_1,b_1}$ and $\epsilon^{PE}_{D_2,b_1}$). This means that workers reduce their effort more strongly when benefits are increased. This effect again increases MH costs.
   (*c*) Because workers find jobs faster, the tax on employed workers to finance the UI scheme does not need to be as high as before. Consumption of employed workers increases and the wedge in the marginal utility from consumption between employed and unemployed workers (and therefore the CS gain) goes up.

(4) There is a new effect due to screening: raising benefits increases offer probabilities (more formally, $\partial g_t/\partial b_1 > 0$ for most t in the screening model, compared to $\partial g_t/\partial b_1 \approx 0$ in the full-information benchmark).[42] Because of its effect on search behaviour, raising b_1 improves the average quality of long-term unemployed applicants. Firms recognise this and are more likely

[42] The reason for $\partial g_t/\partial b_1$ only being approximately rather than exactly zero in the full-information benchmark is that even with perfect information (and no statistical discrimination), benefits still influence the applications-per-vacancy ratio. For example, if benefits were so high that few people search, the applications-per-vacancy ratio would be low and the offer probability would be high for an individual vacancy (since there is no competition). However, we find that $\partial g_t/\partial b_1$ is close to 0 in the full-information benchmark.

Table 8. *Moral Hazard Costs: Components.*

	Component (1)	Full information (2)	Screening (3)	Change (%) (4)	Implied change in the MH cost (%) (5)
1	D_1	11.780	9.584	−18.6	14
2	ϵ_{D_1,b_1} (partial equil.)	0.134	0.160	19.5	–
3	ϵ_{D_1,b_1}	0.131	0.183	39.8	15.4
4	$(b_1+\tau w)/b_1$	1.387	1.318	−4.9	−1.9
5	D_2	11.318	10.622	−6.1	−3.8
6	ϵ_{D_2,b_1} (partial equil.)	0.271	0.296	9.5	–
7	ϵ_{D_2,b_1}	0.256	0.133	−48.2	−29.5
8	$(b_2+\tau w)/b_1$	1.166	1.098	−5.9	−3.6
9	MH	0.468	0.402	−14.1	–

Notes: This table shows each of the components of the moral hazard cost equation (see (7)) for the screening model and the full-information benchmark. The last column shows a simple calculation to show how much changes in each component matter for the overall MH costs: using the equation for the MH cost, the calculation replaces a single component by its value in the screening model (column (3)) while leaving all other components at their value in the full-information benchmark (column (2)). Then, the last column reports how strongly replacing an individual component changes the MH cost (in percentages).

to interview long-term unemployed workers due to their improved beliefs (see Section 4.3.4 below for the mechanisms behind why benefits affect offer rates).

(*a*) The increase in offer rates has a positive externality effect on welfare (EXT >0).

(*b*) In addition, the increase in offer rates also reduces the moral hazard costs, as it makes ϵ_{D_2,b_1} smaller (this offsets the increase in the MH cost from point (3)).

(5) *Taking stock.* Combining points (3) and (4), the moral hazard is 14.1% lower in the screening model (since point (4)(b) more than offsets points (3)(a) and (3)(b)) and the consumption smoothing gain is 6.8% higher. In addition, there is a positive hiring externality, as raising b_1 increases offer rates for the long-term unemployed, so that the externality-adjusted MH costs are 49.7% lower. Therefore, it is optimal to raise b_1.

4.3.3. *Sufficient statistics: more detailed analysis*

Having described the main chain of events, we now turn to a more detailed discussion. We start with the impact of screening on moral hazard costs. Recall the expression for the MH cost:

$$\mathrm{MH}(b_1)=\frac{\overbrace{\epsilon_{D_1,b_1}D_1(\tau+b_1)/b_1+\epsilon_{D_2,b_1}D_2(\tau+b_2)/b_1}^{\mathrm{BR}}}{\underbrace{D_1}_{\mathrm{MRL}}}. \tag{7}$$

The MH cost is the ratio of the behavioural response (the revenue loss that is generated by the elasticities ϵ_{D_1,b_1} and ϵ_{D_2,b_1}) and the mechanical revenue loss from raising b_1, which can be shown to be equal to D_1 for the case of raising b_1. With this equation in mind, we can discuss in detail why moral hazard costs change in the screening model relative to the full-information case. Table 8 shows all components of the moral hazard costs for both the screening model and the full-information benchmark. In the following, we discuss each of the components of the moral hazard equations. When possible, we refer to the point from the chain of events that the component is corresponding to (points (3)(a), (3)(b) and (4)(b) in particular).

- *Average time spent receiving b_1 (point (3)(a)).* Because firms are more likely to hire short-term unemployed individuals, and workers search more in the beginning of the spell, D_1 falls (by 18.6%). This lowers the mechanical revenue loss from raising b_1; that is, there are fewer people who receive b_1, so the mechanical cost of raising b_1 is lower. It also lowers the behavioural response. The logic for the behavioural response is similar: since D_1 is lower, the budgetary impact of raising D_1 by a given percentage (corresponding to ϵ_{D_1,b_1}) is lower. However, the mechanical revenue loss falls more strongly than the behavioural part.[43] This effect alone would lead to higher MH costs, since the MH cost is the ratio between the behavioural and mechanical components. To assess the magnitude of this effect, we perform a back-of-the-envelope calculation where we leave all components of the MH equation at their value from the full-information benchmark and only replace D_1 by its value from the screening model. Substituting in the new D_1 raises the moral hazard costs by 14% (column (5)), which is a significant change.
- *The elasticities.* The elasticities of D_1 and D_2 with respect to benefits (ϵ_{D_1,b_1} and ϵ_{D_2,b_1}) are key components of the moral hazard cost, as they determine the fiscal cost of raising b_1. In our model, these elasticities are determined both by worker and firm behaviour. In the following, we investigate the role of each side of the labour market.

  - *Partial equilibrium (point (3)(b)).* To focus on worker behaviour first, we compute the *partial equilibrium* elasticities (how D_1 and D_2 change when we raise b_1 while holding offer rates constant). We find that these partial equilibrium elasticities increase: $\epsilon^{PE}_{D_1,b_1}$ increases by 19.5% and $\epsilon^{PE}_{D_2,b_1}$ by 9.5%. The interpretation is that workers reduce their search effort more strongly in the screening model when the government raises benefits. Theoretically, it is not obvious how these elasticities should change when we introduce screening and workers anticipate that job-finding rates fall the longer they stay unemployed. Search effort is determined by a forward-looking optimisation problem that does not have an analytical solution. Whether screening increases or reduces the elasticities is therefore a quantitative question.[44]
  - *Including firm responses (point (4)(b)).* When we take firm responses (the reaction of offer rates to benefits) into account, ϵ_{D_1,b_1} is still larger in the screening model. However, there is a substantial change in ϵ_{D_2,b_1}: this elasticity is substantially lower in the screening model than in the full-information benchmark. The reason for this is that in the screening model, raising b_1 increases offer rates for the long-term unemployed. Therefore, raising b_1 does not increase D_2 as much as in the full-information model, because the effect on hiring reduces D_2. In other words, the increase in offer rates at a higher duration ($\partial g_t/\partial b_1 > 0$) partially offsets the decline in search effort ($\partial s_t/\partial b_1 < 0$).
  - Our back-of-the-envelope calculation suggests that these changes in elasticities play a significant role for the moral hazard cost: replacing ϵ_{D_1,b_1} by the value from the screening model (while leaving everything else at the value from the full-information

[43] The reason for this is that the behavioural part also has a second term (relating to b_2), where D_1 is not relevant. Dividing both terms by D_1 leads to the following expression, which is decreasing in D_1: $\text{MH}(b_1) = \epsilon_{D_1,b_1}(\tau + b_1)/b_1 + \epsilon_{D_2,b_1}(D_2/D_1)[(\tau + b_2)/b_1]$.

[44] In the model, search effort has a closed-form solution that depends on the difference between the value function of employment and unemployment ($\Delta V_{j,t+1} = V^e_{t+1} - V^u_{j,t+1}$). Using the closed-form solution for optimal search effort (the formula is given in the Online Appendix), it can be shown that the search elasticities can be written as $\epsilon_{s_t,b_1} = \lambda\epsilon_{\Delta V_{j,t+1},b_1}$, where λ is the exponent of the effort function. Therefore, search elasticities rise when $\epsilon_{\Delta V_{j,t+1},b_1}$ increases. Since the value functions are forward looking, the impact on these elasticities cannot be easily investigated analytically.

benchmark) increases the MH cost by 15.4% and replacing ϵ_{D_2,b_1} lowers the MH cost by 29.5%.

- *The tax.* In the screening model, the tax rate needed to sustain the same schedule that is optimal in the full-information benchmark is lower: the tax rate is only 18.8% rather than 22.8%.[45] Since workers exit unemployment faster, the government does not need to collect as much revenue and can lower the tax rate. From the perspective of the government, the cost of unemployment is $(b_1 + \tau)w$ for the short-term unemployed and $(b_2 + \tau)w$ for the long-term unemployed. This reflects the fact that the government needs to pay benefits and loses tax revenue if workers are unemployed. As the tax rate is lower in the screening model than in the full-information benchmark, this effect reduces the cost of unemployment by making the foregone tax revenue lower and lowers the moral hazard cost. Quantitatively, we find that this plays a smaller role: our calculation suggests that these changes lower the moral hazard costs by between 1.9% (the reduction of $b_1 + \tau$) and 3.6% (the reduction of $b_2 + \tau$). Since the contribution of this effect on the moral hazard cost is small, we have omitted it from the chain of events; note, however, that changes in the tax rate are more significant for the consumption smoothing gain (point (4)(c)).
- *Average time spent receiving b_2 (D_2).* Because of changes in both offer rates and search effort, fewer people enter long-term unemployment and D_2 is 6.1% lower in the screening model. This makes the behavioural response smaller; for a given elasticity, there are fewer people receiving b_2 who reduce their effort. The mechanical revenue loss is not affected, since it depends only on D_1. Therefore, a reduction in D_2 lowers the moral hazard cost, as the behavioural response gets smaller relative to the mechanical revenue loss. However, this effect does not matter much in a quantitative sense, as it only implies a 3.8% reduction in the moral hazard costs according to our back-of-the-envelope calculation. As a result, we have omitted it from the main chain of events.

In addition to the moral hazard costs, the consumption smoothing gain also changes when we introduce screening. Recall that the CS gain is given by

$$\text{CS}(b_1) = \frac{\sum_{t=1}^{D} \beta^t S_t E[u'(c_t^u)]/D_1 - \lambda}{\lambda}.$$

We see that the consumption smoothing gain is higher in the screening model than in the full-information benchmark. This was point (4)(c) in the chain of events. This means that the wedge between the utility of the unemployed and the utility of the employed becomes larger due to screening: it is more valuable to transfer 1€ to the unemployed. Formally, the increase is driven by a fall in the Lagrange multiplier: it falls from 0.0076 (full-information benchmark) to 0.0073 (screening model). Note that the Lagrange multiplier is not as straightforward to interpret as some of the other statistics, as it is the ‘marginal value of public funds’ (the Lagrange multiplier on the government budget constraint) that is jointly determined by all the parameters of the optimisation problem.[46] To see more directly how the gap in marginal utilities between the employed and

[45] Recall that we impose $(b_1, b_2, D) = (0.59, 0.46, 19)$ in the screening model and allow the tax rate to adjust to balance the budget. Therefore, the tax rate differs between the two models.

[46] As discussed in Online Appendix A3, using the FOC for the tax, the Lagrange multiplier can be written as a transformation of the marginal utility of the employed, which also takes the behavioural response to tax changes and effects of the tax on offer rates into account.

Table 9. *Average Marginal Utility of Consumption.*

	Full information	Screening
Employed	0.006	0.006
Unemployed, nineteen or less months	0.011	0.011
Gap (unemployed vs. employed)	73.9%	88.5%

Notes: This table shows the average marginal utility from consumption for the employed and unemployed workers with an unemployment duration of nineteen or less months (which corresponds to the optimal length of the first benefit level in the full-information benchmark). In the third row, the gap in marginal utilities is computed as $(E(u'_u) - E(u'_e))/E(u'_e)$, with $E(u'_u)$ and $E(u'_e)$ being the marginal utilities in unemployment and employment.

unemployed differs, we also compute the average marginal utility in Table 9.[47] The table shows that the gap in marginal utilities is higher in the screening model: the marginal utility of the unemployed is 88.5% higher instead of 73.9%. In the screening model, the tax that is needed to finance the UI schedule is lower since workers find employment more quickly. Therefore, consumption of the employed rises, which lowers their marginal utility.[48] Therefore, the table shows that the gap in marginal utilities rises when we introduce screening, which contributes to the higher consumption smoothing gain.

4.3.4. *Mechanisms behind changes in offer rates*

In the sufficient statistics analysis, we have seen that the hiring externality can lower the effective moral hazard costs of UI reforms (see point (4)(a) in the chain of events). The hiring externality captures the changes in offer probabilities due to a reform (e.g., $\partial g_{jt}/\partial b_1$ in the case of b_1), which has a first-order impact on worker welfare. In this section, we discuss in more detail why implementing the optimal schedule—relative to using the schedule that is optimal in the full-information benchmark—raises offer probabilities. This is useful to better understand how the hiring externality is determined in our model.

Intuitively, our model captures three different reasons why a change in the UI policy can affect offer rates. First, firms optimally form beliefs regarding the productivity of the applicants, depending on the selection of good and bad types over the unemployment spell. For example, if the UI policy keeps good types in the pool for longer, firms will become more optimistic about the productivity of the long-term unemployed and will be more likely to invite a long-term unemployed job seeker for an interview. Second, a firm's interviewing decisions depend on the composition of the pool of applicants. Even if a firm's beliefs are constant, this is an additional channel through which UI reforms can change offer rates. For example, if UI benefits are initially high, this reduces the search effort of the short-term unemployed. As a result, firms will receive fewer applications and are more likely to consider long-term unemployed job seekers (there is a similar effect if the composition of types shifts towards lower-type workers,

[47] Consistent with the formulae for the sufficient statistics, the average marginal utility is computed as $\sum_{t=t_1}^{t_2}(\beta^t S_t)/(\sum_{t=t_1}^{t_2} R^{-t} S_t)E(u'(c_t^u))$ for the unemployed and as $\sum_{t=1}^{T}(\beta^t(1-S_t))/(\sum_{t=1}^{T} R^{-t}(1-S_t))E(u'(c_t^e))$ for the employed.

[48] Note that the average marginal utility of the employed also declines slightly with screening. The reason is that, with screening, workers exit unemployment faster. Thus, the composition of the unemployed shifts towards individuals with a short unemployment duration, who still have more savings available. Changes in savings behaviour can also play a role; for example, when workers anticipate that they will find a job quickly, they consume more of their savings at the beginning of their period of unemployment and their marginal utility declines. From the perspective of the planner, this would make providing benefits less valuable.

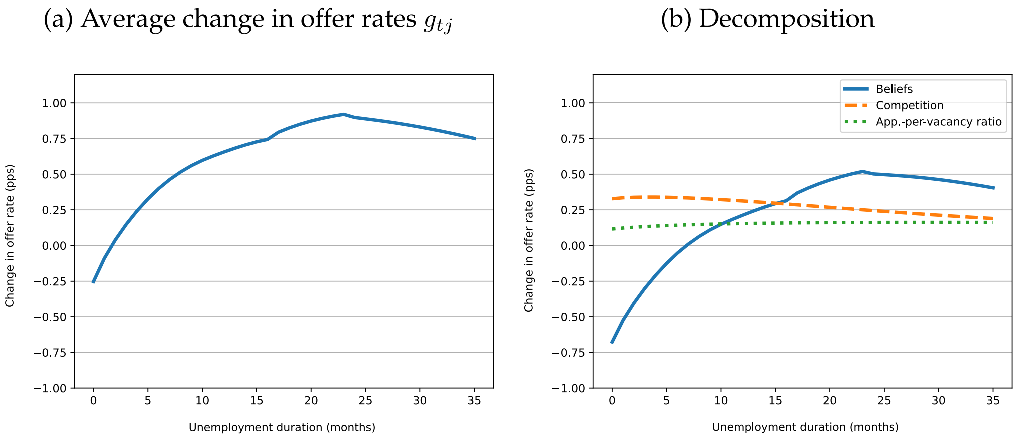

Fig. 8. *The Impact of the UI Policy on Offer Rates: Decomposition.*
Notes: Panel (a) shows the average change in offer rates (in percentage points) for each unemployment duration when comparing the optimal schedule and the schedule that is optimal in the full-information benchmark (both in the screening model). Panel (b) shows the decomposition of this change into the effect of changing beliefs, changing competition and a changing applications-per-vacancy ratio (see the text for details).

who are less likely to be suitable). Third, the UI policy also affects the number of applicants per vacancy. If the overall search effort is so low that firms receive only a few applications per vacancy, they will be more likely to consider the long-term unemployed due to the lack of other applicants.

To illustrate the role of these three channels quantitatively, we perform a decomposition based on the equation for offer probabilities:

$$g_j(t) = \pi_j \int_\phi \exp[-p(\phi, t) \cdot \mu] dN_{1,\sigma}(\phi),$$

$$p(t, \phi) = \sum_{\tilde{t}=1}^{T} \sum_{\tilde{j}=1}^{J} \frac{a_{\tilde{j},\tilde{t}}}{a} \cdot \pi_{\tilde{j}} \cdot P[\Pi(\tilde{\phi}, \tilde{t}) \geq \Pi(\phi, t) \mid \tilde{j}, t, \tilde{t}, \phi].$$

Suppose that we have computed the offer rates in the screening model, given the optimal schedule from the full-information model (g^0_{jt}). Then, we use the screening model to compute offer rates, given the schedule that is actually optimal in the screening model (g^1_{jt}). Panel (a) of Figure 8 first shows the average change in offer probabilities (g^1_{jt} versus g^0_{jt}) from moving to the optimal schedule (relative to the schedule that was optimal in the FI model). As we have seen in previous sections, the hiring externality contributes significantly to the new schedule being optimal. The figure shows that the magnitude of the resulting offer rate changes, indicating that offer rates rise by up to 0.9 percentage points (pps; which is around 6% relative to the pre-reform level).

For the decomposition, the equations above allow us to construct counterfactual offer probabilities, where we only vary certain elements of the equation while keeping the other components fixed. We compute the following three counterfactual offer rates.

- Only $\Pi(\phi, t)$ and $\Pi(\tilde{\phi}, \tilde{t})$ are replaced by the values from g^1_{jt} and everything else stays at its value from g^0_{jt} (*beliefs*): firms only update their beliefs, while the pool of applications they receive is exactly the same.
- Only $a_{\tilde{t},\tilde{k}}/A$ is replaced by the value from g^1_{jt} (*competition by other applicants*): firms keep their old beliefs, but the characteristics of the other applicants change (e.g., more short-term or long-term unemployed).
- Only μ is replaced by the value from g^1_{jt} (*applications per vacancy*): in this case, only the mean number of applicants per vacancy changes.

Note that this is an accounting decomposition rather than a structural one, as the effects are interrelated. The interpretation of the exercise is whether simply substituting in certain factors, such as the updated beliefs, while keeping everything else constant, can explain the changes in offer rates between the model simulations.

Panel (b) of Figure 8 shows the results; that is, to what extent each factor alone can explain the difference in offer rates seen in panel (a). Most strikingly, we see that the change in beliefs plays a significant role in explaining the differences (the solid line). Substituting in the new beliefs, while leaving everything else constant, replicates the shape of the curve in panel (a) and also accounts for a sizeable fraction of the quantitative changes. The changes in beliefs can further explain why offer rates *fall* at the beginning of a spell. As illustrated by (2), beliefs depend on the fraction of suitable applicants at each unemployment duration. Under the optimal schedule, the selection of types over the course of the unemployment spell changes so that job seekers with high unemployment durations are more likely to be suitable. As a result, firms evaluate their productivity more positively and become more likely to interview and subsequently hire them, which slightly hurts unemployment workers with a short unemployment duration (as their offer rate falls slightly).

The second effect that can be evaluated is due to the changing competition from other applicants (the dashed line). This effect increases the likelihood of being hired, as there are fewer individuals with a short unemployment duration and even fewer suitable candidates among those applicants. This in itself increases the probability of being hired at each unemployment duration, and, for example, partially counteracts the negative impact of changing beliefs on offer rates early in the unemployment spell.

Finally, the number of applicants per vacancy declines slightly under the optimal schedule. With fewer applicants per vacancy, it becomes less likely that another applicant will have higher expected productivity. Therefore, from the perspective of a given worker, their probability of being hired increases.

We also performed this exercise separately for the case of varying b_1 and D to see whether the two policy variables have a different impact on offer rates (see Online Appendix Figure A8). We find that the decomposition looks similar in both cases, although the increase in D has a more pronounced effect on offer rates at longer unemployment duration (above two years). In both cases, the belief channel plays the largest role.

Taken together, the increase in offer rates induced by the optimal schedule (which generates the hiring externality and contributes to making this schedule optimal) can be explained by the three effects discussed in this subsection. Most importantly, the decomposition highlights that in a ranking model, offer rates do not depend only on firms' beliefs about the expected productivity of a given applicant, but also on the number and characteristics of the other applicants.

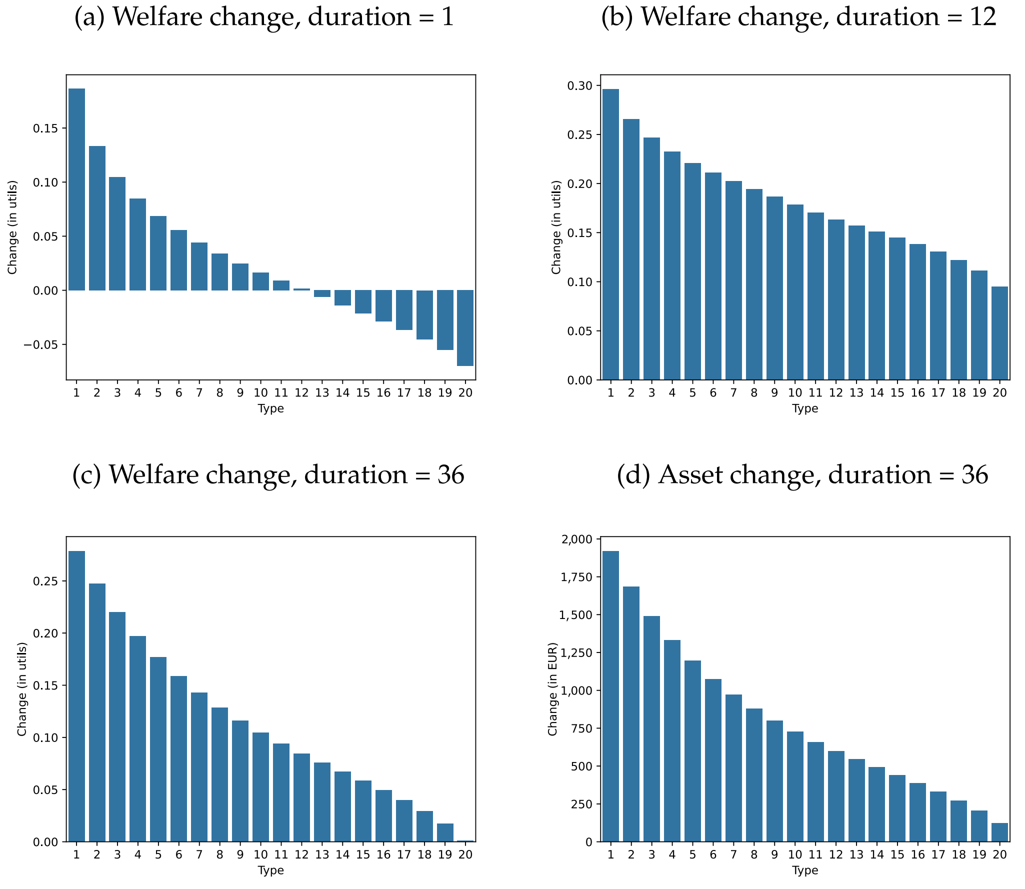

Fig. 9. *Screening Model: Welfare Gains and Losses from the Optimal Schedule.*
Notes: Panels (a)–(c) show the welfare gains and losses for unemployed individuals with different unemployment spell lengths, depending on the productivity type of the individual (averaged over the asset types). These gains and losses are both computed in the screening model, comparing the optimal schedule to the schedule that is optimal in the full-information case. Panel (a) first shows the ex ante welfare change from the perspective of the beginning of the spell and panels (b) and (c) show different durations. Panel (d) shows the mean change in assets later in the spell.

4.4. *Welfare Implications: Who Gains and Loses from Taking Screening into Account?*

To conclude our analysis, we look at the welfare gains from taking screening into account. Who gains and who loses when the government implements the optimal schedule from the screening model, relative to using the schedule that is optimal in the full-information model?

In this context, we discuss two particular cases. The first is using the optimal schedule from the full-information benchmark (with $\sigma = 0$ and without re-estimating the remaining parameters). We impose the schedule that is optimal in this model on the screening model and compare welfare between this schedule and the optimal schedule. Panel (a) of Figure 9 shows the ex ante welfare change. The figure shows that low-type workers gain while high-type workers lose ex ante, with changes being close to 0 in the middle of the distribution. To better assess the magnitude of the welfare changes, we also convert these utility differences to a money metric by computing the asset amount that is required to make the worker indifferent between the two states. The low-type workers gain up to 1,777€ from moving to the optimal schedule, whereas the high-type workers lose up to 937€. Moving to the optimal schedule therefore has important distributional welfare consequences. Note that, on average, the welfare gains and losses somewhat cancel out and the average welfare gain is 259€.

Panels (b) and (c) show how welfare changes for individuals who have been unemployed for longer periods. For example, all workers (irrespective of their type) who have been unemployed for twelve or thirty-six months gain from the transition to the optimal schedule due to its higher generosity and the longer period for which the high level of benefits is paid out. Note that, even unemployed workers at an unemployment duration of thirty-six months are better off with the screening schedule, even though at that point both schedules pay the same level of benefits. This is due to the fact that workers can save part of the higher benefits in previous periods and keep more assets until very late in the spell (panel (d)).

These results complement Table 6 by showing the welfare changes from moving to the optimal schedule, relative to wrongly implementing the optimal schedule from the full-information benchmark. However, note that the full-information benchmark does not match the same data targets as the estimated model, as, for example, job-finding rates are different (recall Figure 5). Therefore, a closely related question is what happens if we re-estimate the full-information model. The interpretation of this second exercise is whether wrongly taking the FI model, estimating it and computing an optimal policy with it provides misleading results. In practice, we find that re-estimating the FI model leads to a fairly similar optimal policy: b_2 and D are identical to the optimal policy in the FIB and b_1 is higher (0.62% rather than 0.59%). Online Appendix Table A3 shows the fit of the full-information model, which deteriorates relative to the full model (as the search and multiple spell moments cannot be fitted as closely as they can with the screening model). Online Appendix Table A4 compares the parameters between the two estimations. The welfare comparisons are quite similar to those shown in Table 9, as the optimal policy is so similar between the FIB and the re-estimated full-information model. The welfare gain from taking screening into account, relative to using the estimated FI model, ranges between 1,258€ for the lowest-type workers and −730€ for the highest-type workers, although it should be noted that the average welfare gains are small and only amount to 134€. Therefore, we find that taking screening into account primarily has a distributional impact on welfare and substantially benefits the lowest-type workers in the population.

5. Conclusion

In this paper, we study how employer screening affects the optimal UI system. We build a model of job search and recruitment behaviour where firms have incomplete information about worker productivity and take the unemployment duration of workers into account when making interviewing decisions. This makes it less likely that the long-term unemployed are invited for interviews. The model is estimated to match several important features of the data regarding job-finding rates, search effort and vacancies.

What are the implications of employers' screening behaviour for the optimal UI system? Our analysis suggests that the presence of screening makes the initial benefit level slightly higher and that benefits should be paid out for a substantially longer time than in the full-information benchmark. We closely examine the reasons for these changes in the optimal policy using a sufficient statistics characterisation of the optimal schedule. Our main finding is that the changes in the optimal schedule in the screening model (relative to the optimal schedule in the full-information benchmark) are due to the following: first, an increase in the insurance gain from providing benefits; second, a decrease in the moral hazard costs; and third, a benefit from the hiring externality, which leads to lower externality-adjusted moral hazard costs. Our analysis shows in detail how these changes in the sufficient statistics are related to workers' changed search incentives and consumption, as well as firms' hiring behaviour. Another key conclusion is that the hiring externality that arises in employer screening models can be quantitatively important for the optimal schedule. This finding complements recent theoretical work about hiring externalities in employer screening models (Lehr, 2017; Kolsrud *et al.*, 2018), and suggests that a better understanding of the hiring externality is important in assessing the welfare gains from changing the UI policy. Finally, we examine the welfare gains and losses when the government takes screening into account and find that the optimal schedule substantially benefits low-type workers, while reducing the welfare of higher-type workers, thereby having a distributional effect on welfare.

Our analysis provides opportunities for further theoretical and empirical research on the relationship between employer screening and UI. On the theoretical front, a particularly important avenue for future research is to endogenise wages. While this is challenging, it would allow the negative informational effect of the unemployment duration to be reflected in both the reduced offer rates and the reduced wages, and lead to a richer analysis of the impact of the UI reforms. On the empirical side, given that we find screening to be important for optimal UI design, it is also important to conduct more empirical work to measure callback and potential offer rates, and to determine how they vary in different economic circumstances (for example, across occupations or the UI benefit regime). This would lead to a more detailed assessment of how the optimal unemployment insurance system should take employer screening into account.

Deutsche Bundesbank, Germany
University of Leicester, CEP & IFS, UK

Additional Supporting Information may be found in the online version of this article:

Online Appendix
Replication Package

References

Alvarez, F., Borovicková, K. and Shimer, R. (2023). 'Decomposing duration dependence in a stopping time model', *The Review of Economic Studies*, doi: 10.1093/restud/rdad109.

Baily, M.N. (1978). 'Some aspects of optimal unemployment insurance', *Journal of Public Economics*, vol. 10(3), pp. 379–402.

Blanchard, O.J. and Diamond, P. (1994). 'Ranking, unemployment duration, and wages', *The Review of Economic Studies*, vol. 61(3), pp. 417–34.

Chetty, R. (2006). 'A general formula for the optimal level of social insurance', *Journal of Public Economics*, vol. 90(10), pp. 1879–901.

Chetty, R. (2008). 'Moral hazard versus liquidity and optimal unemployment insurance', *Journal of Political Economy*, vol. 116(2), pp. 173–234.

DellaVigna, S., Lindner, A., Reizer, B. and Schmieder, J.F. (2017). 'Reference-dependent job search: Evidence from Hungary', *The Quarterly Journal of Economics*, vol. 132(4), pp. 1969–2018.

Doppelt, R. (2016). 'The hazards of unemployment: A macroeconomic model of job search and résumé dynamics', Unpublished Manuscript.

Feng, S., Lefgren, L., Platt, B.C. and Zheng, B. (2019). 'Job search under asymmetric information: Endogenous wage dispersion and unemployment stigma', *Economic Theory*, vol. 67(4), pp. 817–51.

Fernandez-Blanco, J. and Preugschat, E. (2018). 'On the effects of ranking by unemployment duration', *European Economic Review*, vol. 104, pp. 92–110.

Gruber, J. (1997). 'The consumption smoothing benefits of unemployment insurance', *The American Economic Review*, vol. 87(1), pp. 192–205.

Hagedorn, M., Karahan, F., Manovskii, I. and Mitman, K. (2019). 'Unemployment benefits and unemployment in the Great Recession: The role of equilibrium effects', Staff Report 646, Federal Reserve Bank of New York.

Hopenhayn, H.A. and Nicolini, J.P. (1997). 'Optimal unemployment insurance', *Journal of Political Economy*, vol. 105(2), pp. 412–38.

Jarosch, G. and Pilossoph, L. (2019). 'Statistical discrimination and duration dependence in the job finding rate', *The Review of Economic Studies*, vol. 86(4), pp. 1631–65.

Kaas, L. (2010). 'Variable search intensity with coordination unemployment', *The BE Journal of Macroeconomics*, vol. 10(1), doi: 10.2202/1935-1690.2162.

Kollmann, R. (1994). 'The duration of unemployment as a signal', *Economics Letters*, vol. 45(3), pp. 373–77.

Kolsrud, J., Landais, C., Nilsson, P. and Spinnewijn, J. (2018). 'The optimal timing of unemployment benefits: Theory and evidence from Sweden', *American Economic Review*, vol. 108(4–5), pp. 985–1033.

Kroft, K., Lange, F. and Notowidigdo, M.J. (2013). 'Duration dependence and labor market conditions: Evidence from a field experiment', *The Quarterly Journal of Economics*, vol. 128(3), pp. 1123–67.

Krueger, A.B. and Mueller, A. (2011). 'Job search, emotional well-being, and job finding in a period of mass unemployment: Evidence from high-frequency longitudinal data', *Brookings Papers on Economic Activity*, vol. 2011(1), pp. 1–57.

Lalive, R., Landais, C. and Zweimüller, J. (2015). 'Market externalities of large unemployment insurance extension programs', *The American Economic Review*, vol. 105(12), pp. 3564–96.

Landais, C., Michaillat, P. and Saez, E. (2018a). 'A macroeconomic approach to optimal unemployment insurance: Theory', *American Economic Journal: Economic Policy*, vol. 10(2), pp. 152–81.

Landais, C., Michaillat, P. and Saez, E. (2018b). 'A macroeconomic approach to optimal unemployment insurance: Applications', *American Economic Journal: Economic Policy*, vol. 10(2), pp. 182–216.

Lehr, B. (2017). 'Optimal unemployment insurance with endogenous negative duration dependence', *Public Finance Review*, vol. 45(3), pp. 395–422.

Lentz, R. (2009). 'Optimal unemployment insurance in an estimated job search model with savings', *Review of Economic Dynamics*, vol. 12(1), pp. 37–57.

Lichter, A. (2016). 'Benefit duration and job search effort: Evidence from a natural experiment', Discussion Paper 10264, Institute of Labor Economics.

Lise, J. and Robin, J.M. (2017). 'The macrodynamics of sorting between workers and firms', *American Economic Review*, vol. 107(4), pp. 1104–35.

Lockwood, B. (1991). 'Information externalities in the labour market and the duration of unemployment', *The Review of Economic Studies*, vol. 58(4), pp. 733–53.

Marinescu, I. (2017). 'The general equilibrium impacts of unemployment insurance: Evidence from a large online job board', *Journal of Public Economics*, vol. 150, pp. 14–29.

Michaillat, P. (2012). 'Do matching frictions explain unemployment? Not in bad times', *The American Economic Review*, vol. 102(4), pp. 1721–50.

Pavoni, N. (2009). 'Optimal unemployment insurance with human capital depreciation and duration dependence', *International Economic Review*, vol. 50(2), pp. 323–62.

Schmieder, J.F., von Wachter, T. and Bender, S. (2012). 'The effects of extended unemployment insurance over the business cycle: Evidence from regression discontinuity estimates over 20 years', *The Quarterly Journal of Economics*, vol. 127(2), pp. 701–52.

Shimer, R. and Werning, I. (2006). 'On the optimal timing of benefits with heterogeneous workers and human capital depreciation', Working Paper 12230, National Bureau of Economic Research.

Shimer, R. and Werning, I. (2008). 'Liquidity and insurance for the unemployed', *The American Economic Review*, vol. 98(5), pp. 1922–42.

The Economic Journal, 1–41 https://doi.org/10.1093/ej/ueae095
Advance Access Publication Date: 5 November 2024