

Setty, Ofer

Article

Optimal unemployment insurance with monitoring

Quantitative Economics

Provided in Cooperation with:

The Econometric Society

Suggested Citation: Setty, Ofer (2019) : Optimal unemployment insurance with monitoring, Quantitative Economics, ISSN 1759-7331, The Econometric Society, New Haven, CT, Vol. 10, Iss. 2, pp. 693-733,
<https://doi.org/10.3982/QE564>

This Version is available at:

<https://hdl.handle.net/10419/217154>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-nc/4.0/>

Optimal unemployment insurance with monitoring

OFER SETTY

Eitan Berglas School of Economics, Tel Aviv University

I model job-search monitoring in the optimal unemployment insurance framework, in which job-search effort is the worker's private information. In the model, monitoring provides costly information upon which the government conditions unemployment benefits. Using a simple one-period model with two effort levels, I show analytically that the monitoring precision increases and the utility spread decreases if and only if the inverse of the worker's utility in consumption has a convex derivative. The quantitative analysis that follows extends the model by allowing a continuous effort and separations from employment. That analysis highlights two conflicting economic forces affecting the optimal precision of monitoring with respect to the generosity of the welfare system: higher promised utility is associated not only with a higher cost of moral hazard, but also with lower effort and lower value of employment. The result is an inverse U-shaped precision profile with respect to promised utility.

KEYWORDS. Unemployment insurance, optimal contracts, moral hazard, job-search monitoring.

JEL CLASSIFICATION. D82, E24, J64, J65.

1. INTRODUCTION

Most unemployment insurance (UI) programs in the United States include the monitoring of job-search efforts (Grubb (2000)). A typical monitoring policy requires the unemployed worker to record her job-search activities by listing the employers she contacted in a given period. At the employment office, a caseworker evaluates whether the job-search requirements were met by verifying that the contacts are authentic. If the caseworker finds the report unsatisfactory, she may impose sanctions, usually in the form of a reduction in benefits for a limited period.¹

Ofer Setty: ofer.setty@gmail.com

This paper is based on the second chapter of my PhD dissertation. I am grateful for the invaluable guidance of my adviser, Gianluca Violante. I thank the four anonymous referees for numerous useful comments and suggestions. Daniel Bird and Nicola Pavoni were extremely helpful in the development of the theoretical section of this paper. I also thank Yair Antler, Eddie Dekel, Moshe Hazan, Lars Ljungqvist, Zvika Neeman, Ady Pauzner, Dan Simundza, Chris Sleet, and seminar participants at the SED 2013 (Korea), CEMFI, Stockholm University, Wharton Macro workshop, Cornell University, University of Virginia, and Ensaï for very helpful comments. Tali Oberman provided outstanding research assistance.

¹Many countries—including, for example, Australia, Canada, Switzerland, and the United Kingdom—use job-search monitoring for unemployed workers (Grubb (2000)). Since the policy implementation differs across countries, I focus in this paper on job-search monitoring in the United States.

In this paper, I incorporate monitoring into the principal-agent framework of optimal UI, as in [Hopenhayn and Nicolini \(1997\)](#). Monitoring allows the principal (planner) to acquire imperfect information that is related to the job-search effort of the agent (worker). I characterize the optimal contract analytically for a simple one-period model with general preferences. I then calibrate an infinite-horizon model to the US economy and use it to study the properties of the optimal contract with continuous effort relative to an optimal UI model in which monitoring is unavailable.

In optimal UI, a risk-neutral planner insures a risk-averse worker against unemployment by setting transfers during unemployment and a wage tax or a subsidy during employment. During unemployment, the worker searches for a job by exerting effort, the level of which is private information. Since the planner cannot observe the job-search effort, the constant benefits that are implied by the first-best allocation would undermine the worker's incentives to search for a job. Therefore, to solve the incentive-insurance trade-off, benefits should continuously decrease during unemployment, and the wage tax upon reemployment should continuously increase.

I include monitoring in this framework as follows. The planner chooses the quality (precision) of the monitoring technology for the unemployed worker. The cost of monitoring increases with the monitoring's precision, which is correlated with the worker's job-search effort. The planner uses the monitoring signal to improve the efficiency of the contract by conditioning future payments and the wage tax not only on the employment outcome but also on the signal's outcome. These future payments create endogenous sanctions and rewards that, together with the monitoring signal, create effective job-search incentives. By exerting job-search effort, the worker increases the probability of a good signal and, consequently, of higher payments.

I analyze analytically a simple one-period model with binary effort and general preferences from consumption and show analytically that in the optimal contract the signal's precision increases and the spread between future utilities decreases with promised utility if and only if the derivative of the inverse of utility from consumption $(u^{-1})'$ is convex. I show that the driving force of this result is that, while the precision's cost is independent of promised utility, the cost of spreading out future utilities depends on the curvature of $(u^{-1})'$. When $(u^{-1})'$ is convex (concave), the cost of spreading out utilities increases (decreases) with promised utility.

The quantitative analysis uses an extended infinite-horizon model with logarithmic utility from consumption, in which (a) the effort is continuous rather than binary and (b) separations from employment exist. I calibrate the model using information gathered from several micro studies, including the effort distribution of unemployed workers, the probability of sanctions, and statistics regarding the quality of the signal.

The calibrated model shows that depending on the generosity of the welfare system, monitoring bridges up to 50% of the cost associated with moral hazard. The calibrated model also highlights two conflicting economic forces affecting the precision with respect to the generosity of the welfare system. First, for log utility higher promised utility is associated with a higher cost of moral hazard. This economic force, which is also present in the model with binary effort, implies that precision should *increase* with promised utility. Second, higher promised utility is also associated with lower effort and

a lower value of employment. This economic force, which is a consequence of the continuous effort, implies that precision should *decrease* with promised utility. Jointly, the two forces result in an inverse U-shaped curve of monitoring precision with respect to promised utility.

I then extend the one-period model to allow the worker to fake her search behavior. This extension reflects the realistic ability of the worker to control not only her search effort but also the probability of receiving a good signal without actually exerting job-search effort. The extended model is used for two purposes. First, I analyze how the faking behavior affects the optimal contract. In the benchmark model, the cost of spreading out future utilities is limited to the cost associated with compensating the worker for the risk. When faking search is possible, the spread also incentivizes the worker to increase her faking search behavior. Thus, in the presence of faking search effort, the importance of monitoring increases and, therefore, the signal's precision increases and the spread between utilities decreases. Second, I show that the main result of the one-period model without faking search effort, regarding the dynamics of the monitoring signal and the spread with respect to promised utility, is preserved when faking search effort is allowed.

The remainder of the paper is organized as follows. Section 2 surveys the literature on monitoring. Section 3 describes the model. Section 4 provides analytic results using a one-period model with binary effort. Section 5 calibrates and solves quantitatively an infinite-horizon model where effort is continuous and employment is not absorbing. Section 6 extends the one-period model to allow the worker to fake her search behavior. Section 7 concludes.

2. LITERATURE REVIEW

This section first reviews the empirical evidence on the effect of job-search monitoring on labor market outcomes. It then surveys the theoretical literature on monitoring and sanctions.

There is some empirical support for the premise that monitoring is beneficial for reducing the duration of unemployment.² Using the Washington Alternative Work-Search Experiment, which randomly assigns unemployed workers to treatment groups that differ in their job-search requirements, [Johnson and Klepinger \(1994\)](#) found that waiving the weekly requirement to record three contacts increases the average unemployment spell by 3.3 weeks. [Klepinger et al. \(1997\)](#) evaluated the Maryland Unemployment Insurance Work Search Demonstration and found that increasing the number of required contacts from two to four decreases the average unemployment spell by 5.9%. They also find that informing the unemployed workers that the contacts will be verified decreases the average unemployment spell by 7.5%.

The evidence on the effects of sanctions is limited yet in favor of the use of monitoring. In two empirical studies conducted in the Netherlands, [van den Berg, van der](#)

²[van den Berg and van der Klaauw \(2006\)](#) considered a model in which search efficiency is undermined because unemployed workers substitute formal for informal channels. For adverse effects of job search assistance on labor market outcomes, see [van den Berg \(1994\)](#) and [Fougere, Pradel, and Roger \(2009\)](#).

Klaauw, and van Ours (2004) and Abbring, van den Berg, and van Ours (2005) found that the unemployment exit rate doubles following a sanction. Using Swiss data on benefit sanctions, Lalive, van Ours, and Zweimüller (2005) found that both warning about non-compliance with eligibility requirements and enforcement of sanctions for noncompliance increase the unemployment exit rate. In addition, increasing the monitoring intensity reduces the unemployment duration of nonsanctioned workers.

I now review the theoretical literature on monitoring and sanctions. A typical assumption in principal-agent models that entail costly state verification is that monitoring perfectly reveals the worker's hidden information (or action) to the planner. This simplifying assumption rests on Becker's (1968) seminal paper, "Crime and Punishment." In a standard environment, perfect signals allow the planner to get arbitrarily close to the first-best allocation by using a combination of very low monitoring frequencies that cost very little and extremely severe punishments that will never be applied.

Allowing the signal to be imperfect, as I do here, has three salient implications. First, the monitoring precision becomes a choice variable. Second, the contract dictates endogenously limited sanctions and rewards. Third, sanctions are applied in equilibrium. These results are realistic for many applications of monitoring, including that of UI benefits. Specifically, sanctions are used and maximal sanctions are usually not practiced.³ Moreover, it may be infeasible or too costly to perfectly verify the level of the worker's job-search effort.

Since the planner's ability to acquire imperfect information is common to many principal-agent settings, I review models of monitoring in various contexts, with either perfect or imperfect signals. Aiyagari and Alvarez (1995) extended the Atkeson and Lucas (1995) analytical framework by introducing costly monitoring technology. They assume a lower bound on the expected discounted utility that can be assigned to any worker at any date. As in Becker (1968), the monitoring technology is perfect. The solution to their problem, however, differs from Becker's solution because the presence of a lower bound prevents the planner from inflicting Becker's infinite punishment. In their model the investment in monitoring (through the choice of a probability rather than precision) is nonmonotone due to the assumption of a lower bound on utility provided to the worker.

Popov (2009) modeled verification of hidden information as reported by a worker. He keeps the problem nontrivial by assuming that the utility function is bounded from below and that the continuation utility is also bounded. With these assumptions, the contract delivers bounded sanctions and rewards, depending on the verification result. Popov found that monitoring never occurs with certainty and that, for a certain class of utility functions, the principal uses verification regardless of cost.

Ravikumar and Zhang (2012) studied optimal monitoring in a tax compliance context with hidden income in a model with CARA utility. In their model, the later audits are conducted, the more beneficial they are because the likelihood of hidden income increases with time. Since the cost of auditing is constant, the optimal application of monitoring consists of cycles: initially, a low-income taxpayer is unaudited, but with time he faces a positive probability of auditing.

³Grubb (2000) reported an annual frequency of 33.5% or a monthly frequency of 3.3% for the frequency of sanctions due to evidence of active job search.

I now review studies that model monitoring specifically in the context of UI. The model that comes closest to the one I study is by [Boone et al. \(2007\)](#), who analyzed the design of optimal UI in a search equilibrium framework. They allowed the signal to be imperfect, but they restricted the set of policies from which the optimal policy is chosen. First, the planner does not condition benefits on the worker's history; second, the planner can apply only a fixed decrease in benefits for the remainder of the unemployment spell. Hence, there are no additional sanctions (or prizes) that could take place during the remaining unemployment spell. The significance of this difference is that (1) the planner can always further incentivize the agent optimally given all the information that is currently available to her, and (2) the planner can do that while maintaining the lowest possible variation (in terms of utility) in consumption. Their model, however, has the advantage of general equilibrium, which my model lacks.

[Fredriksson and Holmlund \(2006b\)](#) used a search model to compare three different means of improving the efficiency of UI: the duration of benefit payments, monitoring in conjunction with sanctions and workfare. Their analysis suggests that a system with monitoring and sanctions restores search incentives most effectively, since it brings additional incentives for the worker to search actively so as to avoid the sanction.

[Pavoni and Violante \(2007\)](#) considered monitoring as part of an optimal Welfare-to-Work program. In their model, the planner can observe the worker's job-search effort perfectly by paying some cost. As a result, the effort is monitored with certainty, and sanctions or rewards are never needed.

[Fuller, Ravikumar, and Zhang \(2015\)](#) studied concealed earnings, that is, the earnings amassed an unemployed worker becomes employed and still continues to collect benefits. They show that in the optimal contract, the planner monitors the worker at fixed intervals. Similar to my findings, unemployment benefits are relatively flat between verifications but decrease sharply after a verification.

[Fredriksson and Holmlund \(2006a\)](#) surveyed studies on the design of UI in the context of three instruments: time profile, monitoring with sanctions, and workfare. In addition, using a unified theoretical model, they show how these three instruments provide different ways of imposing a penalty on a less active job search.

3. THE MODEL

In this section, I describe my model in detail. It extends the one by [Hopenhayn and Nicolini \(1997\)](#) by introducing costly monitoring, thus expanding the planner's actions space. A simplified one-period version of the model will be studied *analytically* in Section 4. This will be followed by studying *quantitatively* the full-blown version of the model in Section 5. The rest of this section describes the details of the economy.

Preferences: Workers have a period utility $u(c) - a$, where $c \geq 0$ is consumption, a is disutility from job-search effort or work, and u is strictly increasing and strictly concave. The planner is risk-neutral. Both the worker and the planner discount the future at rate β .

Employment and unemployment: The worker is either employed or unemployed. During employment, the worker exerts a constant effort level e_w , with $e_w > 0$, and receives a fixed periodic wage w . Employment ends with probability σ .

During unemployment, the worker exerts an effort level $e \geq 0$ in her job search. This effort is the worker's private information. The job-finding probability $\pi(e)$ is an increasing, differentiable and concave function in e , where $\pi(0) = 0$, and $\lim_{e \rightarrow \infty} \pi(e) \leq 1$.

Monitoring technology: Monitoring provides the planner with a signal on the worker's job-search effort that is either good (g) or bad (b). A meaningless signal, one which does not depend on the worker's effort decision can be thought of as being available at no cost to the planner. However, the planner, at a cost, can affect the mapping from the worker's effort to the good signal's probability. More specific assumptions on the cost structure are provided below.

Information structure: Both the worker and the planner observe the employment state, the monitoring signal, and the work-effort level.⁴ The worker's job-search effort level is her private information. This leads to the moral hazard problem.

Timing: At the beginning of the period, the planner delivers consumption c to the worker. Then, if the worker is unemployed, she chooses an effort level for her job search. If the worker becomes employed, the planner does not monitor her.⁵ If, however, the worker remains unemployed, the planner chooses her investment in monitoring and receives either the good or the bad signal with the probabilities associated with the worker's effort and the monitoring investment.

Possible outcomes: Given the realization of the employment state and the signal, the three possible outcomes are employment (e), unemployment with a good signal (g), and unemployment with a bad signal (b).

4. ANALYTIC ANALYSIS

In this section, I solve a one-period model with a binary effort, $a \in \{0, e\}$. This model, although simple, gives a nontrivial intuition for one of the the main results in the quantitative analysis in Section 5.

In the one-period environment, the planner has to choose transfers to the worker c^x that are conditional on outcome x for $x \in \{e, g, b\}$. The planner's value $V(U)$ is a function of the utility U promised to the worker. In the one-period model, this can also be thought of as the outside option of the worker. The planner's value is the expected wage net of the cost of providing consumption to the worker and the monitoring cost. To lighten notation, I drop the dependency of π on effort for this binary-effort case. Finally, I assume that $\pi \in (0, 1)$.

⁴Assuming the work effort is observable is a standard assumption in the optimal UI literature. Wang and Williamson (2002) considered the case where the worker's effort level affects the probability of transitions both from unemployment to employment and vice versa.

⁵This planner's decision is optimal because employment in the model can be achieved only with a positive effort. Hence, employment perfectly reveals that the worker exerted effort. Notice that the assumption that work cannot be achieved without effort is not a simple normalization.

For the signal structure, I assume that the probability of a good signal is θ when a high effort is exerted and zero when no effort is exerted.⁶ θ is one of the planner's decision variables, and I refer to it as the signal's precision. The planner's cost of monitoring, denoted by $\kappa(\theta)$, is convex and increasing in precision θ , with $\kappa(0)$, the cost of providing a signal that is uninformative, equal to zero, and $\lim_{\theta \rightarrow 1} \kappa(\theta) = \infty$.

The one-period problem is then:

$$\begin{aligned}
 V(U) &= \max_{c^e, c^g, c^b, \theta} \{ \pi(w - c^e) - (1 - \pi)\theta c^g - (1 - \pi)(1 - \theta)c^b - (1 - \pi)\kappa(\theta) \} \\
 \text{s.t. :} & \\
 U &= \pi u(c^e) + (1 - \pi)\theta u(c^g) + (1 - \pi)(1 - \theta)u(c^b) - e, \\
 u(c^b) &\leq \pi u(c^e) + (1 - \pi)\theta u(c^g) + (1 - \pi)(1 - \theta)u(c^b) - e.
 \end{aligned} \tag{1}$$

The first constraint in the problem, called the *Promise-Keeping* (PK) constraint, imposes that the planner delivers in expected terms the utility promised to the worker.

The second constraint, called the *incentive-compatibility* (IC) constraint, requires that the expected utility conditional on effort e will be at least as high as the one without effort (and zero job-finding probability). This constraint incentivizes the worker to search for a job with effort e .⁷ Notice that the term on the LHS, $u(c^b)$, reflects the assumption mentioned above that the probability of a good signal when no effort is exerted is zero.

In what follows, I assume that the wage w is large enough relative to promised utility U to make the effort recommendation optimal.⁸

The following claim determines the ranking of the transfers to the worker.

CLAIM 1. *In the optimal solution, $u(c^e) = u(c^g) > u(c^b) = U$.*

This claim builds on several properties of the problem that are proved in Appendix A: c^b must be lower than at least one of $\{c^e, c^g\}$ to satisfy the IC constraint; c^e has to equal c^g or otherwise the solution can be improved; and the IC constraint is tight.

Since the IC constraint is tight (Lemma 3 in Appendix A), and since the PK is tight as well and has the same right-hand side as the IC, we can substitute $u(c^b)$ in the IC constraint with U and derive $u(c^e) = u(c^g) = U + \frac{e}{\pi + (1 - \pi)\theta}$. Using the values for $\{c^e, c^g, c^b\}$ in the objective function and omitting the term πw , which is independent of the choice

⁶In Section 6, I extend the model to allow for a good signal also when no search effort is exerted.

⁷For sufficiently high promised utility, creating incentives by spreading future promised utilities is too costly; hence, the planner recommends low job-search effort and implements full insurance (which Pavoni and Violante (2007) refer to as "social Assistance"). To fully characterize the optimal monitoring policy, I describe the monitoring policy while assuming that it is always desirable to create incentives for the worker to expend a high job-search effort. Notice that since I abstract from allowing the principal to choose zero effort, I also abstract from lotteries over payoffs. According to the calibration in the next section, social assistance is optimal only for those values of promised utility associated with consumption levels that are more than 10 times the government's balanced-budget point.

⁸Otherwise, if the planner recommends a zero effort, then moral hazard is no longer present in the problem and the planner perfectly insures the worker.

variables, leads to the following convex optimization problem, whose solution for θ is identical to that of Problem 1:

$$V = \max_{\theta} \left\{ \alpha(\theta) u^{-1} \left(U + \frac{e}{\alpha(\theta)} \right) + (1 - \alpha(\theta)) u^{-1}(U) + (1 - \pi) \kappa(\theta) \right\}, \quad (2)$$

where $\alpha \equiv \pi + (1 - \pi)\theta$.

To understand the role of monitoring precision in this problem, consider the solution to the first best. In the first best, the planner observes the worker's effort, so no monitoring is required. The first-best allocation is then a fixed transfer (independent of the employment outcome) that is equal to $u^{-1}(U + e)$. In this case, the planner compensates the worker only for her effort.

The planner's value in the constrained problem, Problem 2, differs from that in the first best in two aspects. First, monitoring may be used upon unemployment with a cost of $(1 - \pi)\kappa(\theta)$. Second, the planner is required to create a spread in consumption conditional on outcomes. Therefore, in Problem 2, the planner delivers utility to the worker as a lottery between $u^{-1}(U + \frac{e}{\alpha})$ with probability α and $u^{-1}(U)$ with probability $(1 - \alpha)$. I refer to the difference between the two utilities $(U + \frac{e}{\alpha}, U)$ as the *spread*.

The average utility delivered through the lottery is $U + e$. This is, by construction, equal to the utility delivered in the first best. Therefore, the only role of the signal in this problem is to reduce the risk associated with the spread and thus reduce the cost of delivering utility as a lottery rather than as a certain outcome. Indeed, if the signal were without cost, the planner would set $\theta = 1$, and both the allocation and the planner's cost would be identical to those of the first best. (To see this, substitute $\theta = 1$ and $\kappa(\theta) = 0$ in Problem 2 and get the first-best value.)

Proposition 1 characterizes the contract with respect to promised utility U , depending on the concavity of $(u^{-1})'(\cdot)$.

PROPOSITION 1. *If $(u^{-1})'(\cdot)$ is convex, the solution to Problem 2 has the following characteristics:*

- (i) *the optimal signal's precision (θ) increases with the outside option (U);*
- (ii) *the utility spread $(\frac{e}{\pi + (1 - \pi)\theta})$ decreases with promised utility (U);*
- (iii) *the cost of spreading out utility increases with promised utility (U);*
- (iv) *the converse version of (i)–(iii) holds when $(u^{-1})'(\cdot)$ is concave.*

For any level of promised utility, the planner weighs the cost of the signal against its benefit of reducing the risk associated with the spread. The signal's cost does not depend on promised utility. When $(u^{-1})'(\cdot)$ is convex, the cost of spreading out utilities increases with promised utility (see Appendix A for the proof of Proposition 1 (iii)). When this condition holds, the value of monitoring increases, and the planner increases her investment in the signal. This, in turn, results in a smaller spread between utilities.

Examples of utility functions that satisfy a convex $(u^{-1})'(\cdot)$ are Increasing Absolute Risk Aversion (IARA), Constant Absolute Risk Aversion (CARA), and Constant Relative Risk Aversion (CRRA) with a coefficient of relative risk aversion of at least 0.5, for which log utility belongs.

5. QUANTITATIVE ANALYSIS

I now turn to a quantitative analysis of the infinite-horizon model with continuous job-search effort and separations from employment as described in Section 3.

5.1 *Signal structure*

I start by making explicit assumptions on the signal structure. The signal is a mapping from any effort, and, in particular, the recommended/required effort level, to the probability of a good signal. I assume a simple, yet flexible, structure that allows matching the available data on the probability of sanctions and on statistics regarding the quality of the signal, given the distribution of effort of unemployed workers. In doing so, the main assumption that I make is that the probability of a good signal is linear in effort.⁹ With that assumption, the signal's structure requires only two parameters: the slope of the signal with respect to the effort and the probability of a good signal at the recommended effort.

An example of that signal structure is provided in Figure 1. The horizontal axis in this figure is the support for a job-search effort ranging from zero to 1.5 (which is the value calibrated below for the effort that guarantees finding a job). The vertical axis is the probability of a good signal. The dashed line indicates the recommended effort level (at an effort level of 0.2 in this example). The solid line is an example of the mapping between effort and the probability of a good signal. Hence, it is the measure of workers at each effort level who receives a positive signal. In this particular example, there are two properties: (1) zero effort is associated with a positive probability for a good signal, and (2) there is a threshold effort level (at an effort of 1.2) from which the probability of

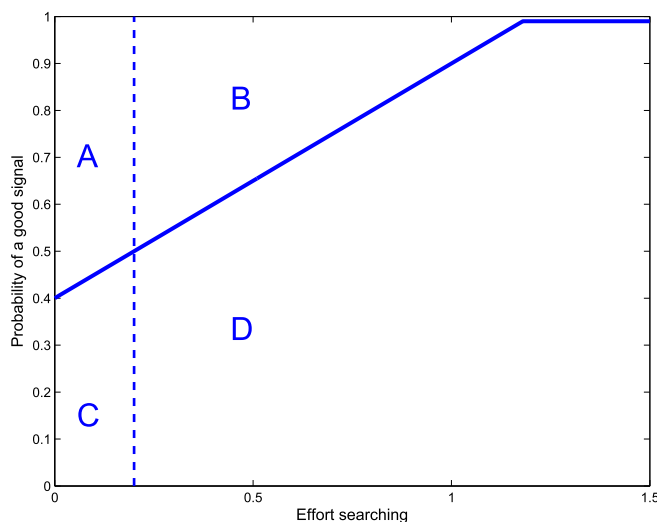


FIGURE 1. Signal structure.

⁹Relaxing that assumption by adding more parameters has little effect on the analysis.

a good signal onwards is 1.0.¹⁰ Notice that those two properties are not imposed by the general signal's structure.¹¹

The slope of the signal determines to what extent the signal is revealing. In particular, the higher the slope, the more informative the signal: a flat function carries no information, while a dramatically increasing one allows the planner to distinguish clearly between the recommended effort and other effort levels.

The two lines—the probability of a good signal conditional on effort and the vertical line representing the required job-search effort—divide the graph into four distinct regions, denoted by A, B, C, and D.

In regions A and C to the left of the vertical line, workers have not exerted sufficient effort. In regions A and B above the solid line, (the signal) workers are sanctioned. Taken together, the regions can be characterized as follows. Region A represents just (or fair) sanctions; it comprises those workers who did not exert the required effort and were sanctioned. Region B represents *underpayments* and comprises those workers who exerted the required effort (or higher) but were unjustly sanctioned. In contrast, region C represents *overpayments* and comprises those workers who did not exert the required effort but still received benefits. Finally, region D comprises workers who exerted the required effort and received benefits. This mapping of the distribution of workers across those four regions between the model and the data will play a key role in determining the signal's parameters in the calibration.

Given the assumptions above about the signal's structure, the functional form for the probability of a good signal given the effort recommendation \underline{e} and the worker's actual effort e is given by $\theta(\underline{e}, e) = \min(1, \max(0, \underline{\theta} + \zeta \cdot (e - \underline{e}))$, where $\underline{\theta}$ is the good signal's probability if the worker follows the planner's recommendation \underline{e} and ζ is the slope of the signal with respect to deviations from the recommended signal.¹²

The slope ζ is the key parameter for determining the precision of the monitoring technology. If, for example, the probability of the signal is invariant to the effort, the signal is worthless as it provides no information about the worker's effort. The sharper the response of the signal to the worker's effort, the sharper its precision. I therefore use this parameter to determine the quality, or precision, of the signal.¹³ The cost of monitoring, which is denoted by $\hat{\kappa}(\zeta)$, is assumed to be increasing in the slope.

¹⁰The signal's structure in the quantitative analysis is richer than the one in Section 3 in that it allows for a good signal even without positive job-search effort.

¹¹Other parameters could imply a zero probability for a good signal for effort levels up to some threshold, or a probability lower than one for even the maximum effort level.

¹²Notice that a worker who exerts an effort higher than the recommended one receives a good signal with a (weakly) higher probability. An alternative modeling choice would be to set the probability of a good signal to be decreasing in the distance between the recommended effort and the actual one. Numerically, this does not matter as the deviation that is binding is downwards, not upwards. Moreover, it is atypical to sanction workers for exerting an effort level that is too *high*.

¹³The other key parameter is $\underline{\theta}$ —the probability of a good signal at the recommended effort. I focus on ζ and note that the main results in this paper do not depend on which margin of monitoring's quality the planner can control.

5.2 The planner's problems

The optimal contract between the planner and the worker requires that the benefits and the wage tax be conditioned on the worker's entire history. Spear and Srivastava (1987) showed that all the relevant information for the recursive contract can be contained in a one-dimensional object. In the monitoring contract, as in the UI contract, this one-dimensional state is the expected discounted utility U promised to the worker at the beginning of each period.¹⁴ This value is then updated at the end of each period, according to the outcomes. Hence, this state (U) is governed by all of the relevant information in the worker's history. Although this state is not a primitive of the model, using it makes the problem tractable. Once the model is solved, the state can be used to recover the allocation for each type of worker. I maintain the standard assumption that the planner is able to fully commit to the contract.

In what follows, I present the planner's problems for an employed worker and for an unemployed one.

5.2.1 The planner's problem for an employed worker Let $W(U)$ be the planner's value from an employed worker who has promised utility U . Similarly, let $V(U)$ be the planner's value from an unemployed worker who has promised utility U .¹⁵ The planner's problem for an employed worker is

$$\begin{aligned} W(U) &= \max_{c, U^e, U^u} -c + w + \beta(1 - \sigma)W(U^e) + \beta\sigma V(U^u) \\ \text{s.t. :} & \\ U &= u(c) - e_w + \beta(1 - \sigma)U^e + \beta\sigma U^u, \end{aligned} \tag{3}$$

where U^e and U^u are the future promised utility levels contingent on employment and unemployment, respectively. If $c > w$, the planner delivers the difference to the worker as a wage subsidy; if $c < w$, the planner extracts the difference as a wage tax. The PK constraint imposes that the planner delivers in expected terms the utility promised to the worker.

5.2.2 The planner's problem for an unemployed worker In the problem for the unemployed worker, the planner has six decision variables: consumption c ; three continuation values, one for each possible outcome: employment U^e , unemployment with a good signal U^g , and unemployment with a bad signal U^b ; an effort recommendation e ; and the monitoring precision θ .

The job-search effort recommendation needs to be supported by appropriate incentives. This is achieved with the IC constraint, which guarantees that the expected utility

¹⁴The initial promised utility is given exogenously. It is easy to see that in this setting where the planner has a cost of providing the utility for the worker, the planner would choose to provide the worker with the minimum promised utility.

¹⁵The planner's problems can be formulated in two alternative approaches. The first is maximizing welfare subject to a resource constraint. The second is maximizing the revenue from a contract that provides the worker with some utility level. My analysis follows the *optimal unemployment insurance* literature that traditionally uses the second approach.

for a worker who exerts the recommended job-search effort maximizes the value of the worker, given the payoffs and the monitoring precision.

Let $V(U)$ be the planner's value from an unemployed worker who has promised utility U . The planner's problem for that worker is

$$\begin{aligned}
 V(U) = & \max_{c, U^e, U^g, U^b, e, \zeta} \{-c + \beta \pi(e) W(U^e) \\
 & + (1 - \pi(e)) [\underline{\theta} V(U^g) + \beta(1 - \underline{\theta}) V(U^b) - \hat{\kappa}(\zeta)]\} \\
 \text{s.t. :} & \\
 U = & u(c) - e + \beta \pi(e) U^e + (1 - \pi(e)) [\underline{\theta} U^g + (1 - \underline{\theta}) U^b], \\
 e \in & \operatorname{argmax}_{e \in [0, \infty)} \{u(c) - e + \beta \pi(e) U^e + \beta(1 - \pi(e)) [\theta(\underline{e}, e) U^g + (1 - \theta(\underline{e}, e)) U^b]\},
 \end{aligned} \tag{4}$$

where $\theta(\underline{e}, e) = \min(1, \max(0, \underline{\theta} + \zeta \cdot (e - \underline{e}))$ as defined above.

The objective function comprises the cost of consumption payments to the worker and the discounted weighted values of the three possible outcomes. The first constraint (PK) takes into account the three possible outcomes and their associated probabilities. The second constraint is the IC constraint described above.

5.3 Calibration

In this subsection, I calibrate the model to US data. Some parameters are determined exogenously to the model while several key parameters are determined endogenously.

The unit of time is set to one month. Preferences are log utility in consumption. The monthly discount factor β is set to 0.9975 to match an annual interest rate of 3%.¹⁶ Monthly earnings, w , are set to \$4000, which is approximately the mean monthly earnings of all workers (DOL (2015)).

We are left with several key parameters to be determined. Those are the utility cost of effort (both for job search and for work) and the signal structure. The external calibration of the monetary cost of monitoring, $\hat{\kappa}$, concludes this subsection.

5.3.1 Cost of effort Recall that the worker's utility function is $\log(c) - a$, where c is consumption and a is disutility from either searching or working. While I allow a continuous choice of job-search effort, I assume only one level of effort for employed workers (e_w). This is done for simplicity and to focus on the effort choice during unemployment.

The utility cost of working is calibrated on the basis of micro evidence regarding the cost of labor participation in consumption-equivalent terms. Point estimates for that cost vary between 20% and 60%.¹⁷ In what follows, I use the middle of the range (40%)

¹⁶I use Moody's Seasoned Aaa as the basis for the interest rate (Source: Board of Governors of the Federal Reserve System (US), Release: H.15 Selected Interest Rates). This is a nominal interest rate that needs to be adjusted for inflation. For inflation, I use the consumer price index for all urban consumers: all items (Source: US Bureau of Labor Statistics, Release: Consumer Price Index), forward-looking inflation. The average for years 2000–2014 is 3%, and is quite insensitive to the sample period.

¹⁷Within labor force participation models, Attanasio, Low, and Sanchez-Marcos (2008), Hausman (1980), Cogan (1981), and Eckstein and Wolpin (1989) computed costs of, respectively, 21%, 27%, 41%, and 62% in consumption-equivalent terms.

to determine the parametric value of e_w . The static condition for participation is given by the following condition:

$$\log(c^P) - e_w = \log(c^{NP}), \quad (5)$$

where c^P and c^{NP} are consumption levels for participants and nonparticipants, respectively. Given the point estimate of 40% as the cost of participation, we get from equation (5) that $e_w = 0.51$.

The utility cost of job search, e , which is a mapping between effort and a job finding probability, is disciplined by various sources of micro evidence. The cost function is borrowed from [Paserman \(2008\)](#), who uses National Longitudinal Survey of Youth data to estimate the following relationship between the job-search utility cost e and the job-finding probability π :

$$e(\pi) = k\pi^{1+\eta}, \quad (6)$$

where k and η are parameters. η , which represents the curvature of the cost function, is a model-invariant parameter taken directly from Paserman's estimation to be 0.4. k , which is model specific, represents the scale of the cost function, telling us how costly it would be to obtain a job with probability one. Using Paserman's functional form, I infer k from the job-finding probability in the data as follows.

An unemployed worker chooses job-search effort e to maximize the following dynamic problem:

$$U = \max_e \{ \log(b \cdot w) - e + \beta\pi(e)E + \beta(1 - \pi(e))U \},$$

where U and E are the values of being unemployed and employed, respectively, and b is the replacement rate that multiplies the wage w .¹⁸ The first-order condition of this problem is

$$\pi'(e) = \frac{1}{\beta(E - U)}. \quad (7)$$

The planner's value of an employed worker, taking into account the separation rate, is

$$E = \log(w) - e_w + \beta(1 - \sigma)E + \beta\sigma U. \quad (8)$$

Notice that since there are no decisions to be made during employment, the optimal value of employment for the planner is given directly. By using e^* to denote the optimal job-search effort and π^* to denote the job-finding probability associated with e^* , by definition we have

$$U = \log(b \cdot w) - e^* + \beta\pi^*E + \beta(1 - \pi^*)U. \quad (9)$$

¹⁸My model also has monitoring and, therefore, the future outcomes should include good and bad signals as well. That full analysis (available upon request) delivers almost the same parameters as the one that follows. For simplicity, I use this simpler specification as an approximation.

Combining equations (8) and (9) give

$$E - U = \frac{\log\left(\frac{1}{b}\right) - e_w + e^*}{1 - \beta[(1 - \sigma) - \pi^*]}. \quad (10)$$

Using equation (7) together with equation (10), we arrive at the following condition:

$$\pi'^* = \frac{1 - \beta[(1 - \sigma) - \pi^*]}{\beta\left(\log\left(\frac{1}{b}\right) - e_w + e^*\right)}. \quad (11)$$

Using Paserman's functional form for π and π' , we arrive at the following two equations in two variables (e, k):

$$\frac{1}{1 + \eta} \frac{(e^*)^{\frac{1}{1+\eta}-1}}{k^{\frac{1}{1+\eta}}} = \frac{1 - \beta[(1 - \sigma) - \pi^*]}{\beta\left(\log\left(\frac{1}{b}\right) - e_w - e^*\right)}, \quad (12)$$

$$\pi^* = \left(\frac{e^*}{k}\right)^{\frac{1}{1+\eta}}. \quad (13)$$

While there is a distribution of effort levels in the data (which I will use later on in this section) the prime purpose of the analysis in this subsection is the parameter k that determines the level cost of search. I therefore use equations (12) and (13) for the average worker, with the average job-search effort \bar{e} (associated with the average job-finding probability $\bar{\pi}$), and the average separation rate $\bar{\sigma}$.

The values for η , β , and e_w are given above. The replacement rate of benefits, $b = \frac{1}{2}$ is based on the typical replacement rate in the United States. Finally, the mean job-finding rate $\bar{\pi}$ and the mean (monthly) separation rate $\bar{\sigma}$ are equal to 0.40 and 0.02, respectively, based on Current population survey (CPS) data.¹⁹

The solution to the system of equations (12) and (13) is $\bar{e} = 0.45$ and $k = 1.50$. This implies that the disutility associated with finding a job with the mean job-finding probability ($\bar{e} = 0.45$) is slightly lower than that of a full-time job ($e_w = 0.51$), and the disutility associated with finding a job with certainty is considerably higher.

5.3.2 Signal structure Figure 1, which was introduced earlier in this section, is useful for matching the data on sanctions, overpayments, and underpayments. However, before using these data, we need to know the effort distribution in the data, as there is no reason to assume that workers are uniformly distributed over the effort's support.

Using data from the American Time Use Surveys (ATUS) from 2003 to 2007, [Krueger and Mueller \(2010\)](#) studied the search behavior of unemployed workers. Figure 2, which is taken from their paper (their Figure 2, p. 301), provides the kernel density of job search in minutes per day conditional on nonzero search. Since the extensive margin for search

¹⁹I use CPS data for individuals who are 25 years and older, 1994:01–2012:07, and the same procedure described in [Elsby, Hobijn, and Şahin \(2010\)](#) and [Shimer \(2012\)](#) to correct for short-term unemployment.

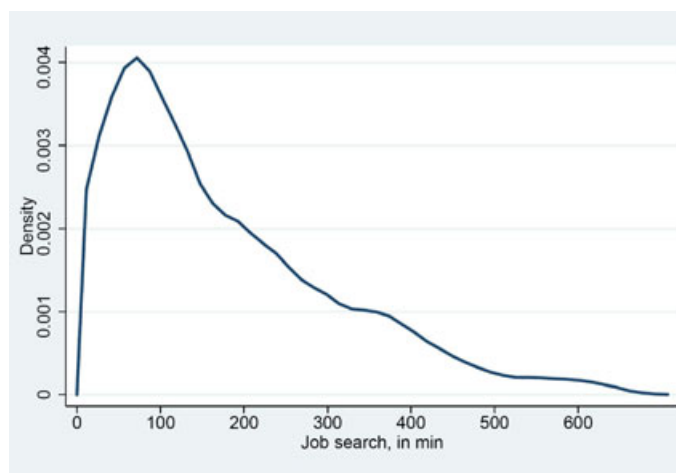


FIGURE 2. Job Search in Minutes per Day (from Krueger and Mueller (2010)).

(for UI-eligible workers) is 27.9%, this distribution reflects the search effort only of workers who actively search.

While this evidence on the time spent searching is useful, it is only based on one day of reporting. To construct the *monthly* distribution of search time, some additional assumptions must be made with respect to the extensive and intensive distribution of search over days. Notice that there is a substantial flexibility in these assumptions. On one extreme, one could assume that the unemployed workers do not rotate in either the extensive or the intensive search effort. In other words, only 27.9% of UI-eligible workers ever search, and when they do, they search for the same amount of time every day. This assumption seems to be unlikely.²⁰

At the other extreme, one could assume that unemployed workers rotate on both the extensive and intensive margins in the ATUS data. This would mean that after aggregation at the monthly level, all workers exert the same search effort. This assumption seems unlikely as well.

I make an intermediate assumption: that workers rotate on the extensive margin but not on the intensive margin. I then assume that effort is linear in the time spent searching and that the maximum effort, the one that guarantees a job is associated with the right tail of the job-search duration at about 600 minutes per day. With those assumptions, the distribution of effort $e \in [0, k]$ is identical to the distribution of time spent searching in Figure 2.

Next, I present evidence on the probability of a sanction. Grubb (2000) used data from national authorities in the United States to calculate the frequency of sanctions imposed for various reasons. For sanctions due to *evidence of active job search*, he reports an annual frequency of 33.5% or a monthly frequency of 3.3%.

²⁰There are two pieces of information that are inconsistent with this assumption. First, in Germany, the only country with two consecutive diary days, 7% of those who did not search in the previous day search on the second day (see Krueger and Mueller (2012), p. 777). Second, the monthly probability of a sanction of 3.3% (Grubb (2000)) is much lower than the probability of monitoring.

The third and final data source that I use for calibration of the signal's structure is measures for overpayments and underpayments. Woodbury (2002) uses data from two programs in the United States that are administered by the Labor Department: the *Benefit Accuracy Measurement*, which studies overpayments of UI benefits, and the *Denied Claims Accuracy*, which tracks the extent to which UI claims are incorrectly (or wrongfully) denied and, therefore, result in underpayments.

Using this information, which reveals, at least partially, to what extent the monitoring technology used by caseworkers is precise, Woodbury reported (Table 1 in his paper) an overpayment rate of 7.2%. The fraction of overpayments due to nonseparation errors out of total overpayments is 19.8% (DOL (2006)). Therefore, the overpayment rate due to nonseparation errors is 1.4% ($19.8\% \cdot 7.2\%$). Meanwhile, the total fraction of underpayments according to Woodbury is 3.4%, of which 57% (Vroman and Woodbury (2001)) are due to nonseparation errors, resulting in nonseparation underpayments of 1.9%.

Using those pieces of evidence, we have altogether three moments to match: the fraction of sanctioned workers, the fraction of overpayments, and the fraction of underpayments. There is an equivalent number of parameters that need to be estimated using the data: the minimum job-search effort required, denoted by \underline{e} , and the two parameters of the signal: the probability of a good signal given the recommended effort and the slope of the signal with respect to the effort.

Taking into account the effort distribution of the workers as explained above, the parameters that minimize the distance between the model and data moments are as follows. The probability of a good signal given the recommended effort is 0.6. Therefore, there is a relatively high chance of sanctions at the lowest effort required. The minimum effort requirement is low at $e = 0.02$ and the slope of the signal with respect to the effort is 7.²¹ Explicitly, the probability of a good signal as a function of effort e , given the recommended effort \underline{e} , is $\theta(e, \underline{e}) = \min(1, \max(0, 0.6 + 7 \cdot (e - \underline{e})))$.

5.3.3 Monitoring cost The cost of monitoring has been estimated in several contexts. Ashenfelter, Ashmore, and Deschenes (2005) reported that, in the experiments they evaluated, the additional weekly processing costs per claim associated with monitoring varied from \$1 to \$15. These costs were mainly due to the added staff time required to go through the supplemental eligibility checks and to monitor search effort.

Corson and Nicholson (1985) and Meyer (1995) evaluated the Charleston Experiment, which sought to strengthen the monitoring of UI work test, offer job-search workshops to job seekers, and enhance the placement of those job seekers through additional services. In this experiment, UI claimants were divided into three groups differing in the intensity of the treatment. Group 3 was only subject to additional eligibility checks. Corson and Nicholson estimated that the program cost per claimant in this group was roughly \$9 per week. For this same experiment, Meyer reported smaller weekly costs, around \$6, because he measured costs for the treatment group net those for the control

²¹The low minimum requirement is expected given, on the one hand, the large mass of unemployed workers with low search effort (Figure 2) and, on the other hand, the low probability of both sanctions of 3.3% (implying that regions A and B are small) and overpayments (region C).

group, which should be interpreted as the costs of administering UI and, as such, should be excluded from the calculation.

In data from the Minnesota Family Investment Program (2000), each caseworker was responsible for 100 clients and, among other tasks, was assigned to apply sanctions, assist with housing, and document client activities. I obtain a monthly cost of \$30 per unemployed worker monitored.²² This value is an upper bound since the caseworker is involved in more activities than monitoring alone.

In sum, there is a wide range of estimates for the monitoring cost, varying between \$4 and \$60 per worker per month. I choose the cost of \$5, at the low end of available estimates, for expositional purposes, because at that cost the results are most stark. To complement the analysis with the low estimate I report, at the end of this section, the results for costs up to \$20, for which the results are qualitatively the same and quantitatively similar.

The estimation of the curvature of the cost function with respect to the slope, ζ , is challenging as it requires data on different monitoring technologies as well as how workers respond to such changes by altering their job-search effort. Instead of estimating that curve, I use a simple linear cost curve, in which a slope of zero is costless (and carries no information) and the calibrated slope is provided at the calibrated monetary cost.

5.4 Results

In this subsection, I describe the results of the quantitative analysis regarding the planner's optimal choices. The analysis is done in stages to allow for a good understanding of the economic forces that shape the planner's decision. I start with the case of two effort levels and then move to a continuous effort. Within each of those analyses, I first study the planner's decisions for an exogenously given free monitoring technology. This type of analysis sheds light on how monitoring affects the planner's value. I then allow the planner to choose the optimal (costly) precision level as well.

All the analyses are done based on the calibration above. For the simplified cases where I study only two effort levels, I set the positive job-search effort to be the same as that for employment. This is done for simplicity, as the main results are valid for any job-search effort.

I then analyze the implications of the choices for consumption dynamics over the unemployment spell. I conclude this subsection by discussing on the sensitivity of the results to the cost of monitoring. Separations are allowed throughout the analysis in this section.

5.4.1 Two effort levels with exogenous monitoring Figure 3 presents the effect of monitoring on the planner's value for the simple case, in which there are only two effort levels and the precision is exogenous. The primary horizontal axis is the single state variable of the problem: promised utility. The secondary horizontal axis presents the certainty-equivalent consumption levels for the promised utility axis (assuming a constant disutility from work). The wide span of consumption, corresponding to a large span in the

²²Using an annual salary of \$36,000 as the employment cost of the caseworker and dividing it by the number of cases she handled

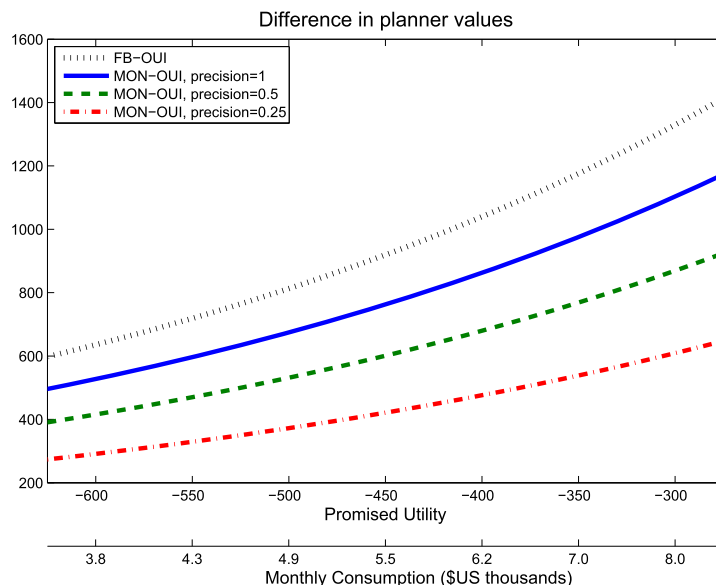


FIGURE 3. Planner values for two effort levels and exogenous precision.

generosity of the program, allows a complete characterization of the planner's choices.²³ The vertical axis shows the difference in value for the planner from a shift from optimal unemployment insurance, where monitoring is absent, to either the first best or monitoring.

The dotted line shows the difference in the planner's value between the first-best allocation, absent moral hazard, and optimal UI. The vertical axis units are simply \$US, as the planner is risk neutral. It is clear from this figure, and consistent with the analytic analysis above, that for log utility the cost of moral hazard, that is, how much the planner can improve her value in the absence of moral hazard—increases with promised utility.

The solid line shows the difference between the value of the problem with monitoring and the value of the problem without monitoring (optimal UI) for maximum precision. The closeness of the solid line to the dotted line indicates that with the maximum precision level, monitoring is effective in bridging over most of the moral hazard cost.²⁴

The next two lines—the dashed line and the dash-dotted line—show the difference between the value of the problem with monitoring and the value of optimal UI for precision levels of 0.5 and 0.25, respectively. As the quality of the signal decreases, so too does the planner's ability to improve her value relative to optimal unemployment insurance.

²³One particular interesting level of generosity is the one implied by the “balanced budget point,” at which the planner's revenues from taxes exactly equal to her cost of benefits. The generosity level associated with providing constant consumption equal to the wage of \$4000 is calculated as follows: $U_0 = \frac{u(c) - \epsilon w}{1 - \beta}$. In the quantitative analysis, one unit of consumption represents \$10,000. Therefore, plugging-in $c = w = 0.4$ in that formula gives $U_0 = -570$.

²⁴It is possible to arbitrarily increase the maximum precision to ensure that monitoring will get arbitrarily close to the first-best. Such a monitoring technology, however, seems unrealistic.

It is evident from Figure 3 that the increased importance of information holds for both a higher precision signal and the removal of moral hazard. This motivates measures that specify the value of monitoring *relative* to the first best. One such measure, which I simply call “savings” and denote by s , is $\frac{V^{\text{MON}} - V^{\text{OUI}}}{V^{\text{FB}} - V^{\text{OUI}}}$. The denominator is the difference between the planner’s value under the first best and her value under optimal UI. This denominator represents potential savings, as the first best cannot be improved upon. The numerator is the actual savings when moving from optimal UI to monitoring. Since (optimal) monitoring cannot be worse than optimal UI and cannot do better than the first best, the savings’ support is $[0, 1]$.

The top panel of Figure 4 shows that metric (savings) for the three cases of monitoring shown in Figure 3. While all lines in Figure 3 are increasing with promised utility, the level of savings is constant. To contrast the savings with its source, the additional information, the middle panel of Figure 4 shows the (exogenous, for now) precision of monitoring.

Figure 5 shows the channel through which the information used in monitoring improves the value of the planner. Focusing on a small interval of promised utility levels around the balanced-budget point, the figure shows the mapping between current promised utility and future promised utility, conditional on outcomes, for both the optimal UI policy (with solid lines) and monitoring with precision of one (with dashed lines). In UI, the information that can be used to condition future outcomes is limited to the employment outcome. As a consequence, a relatively large wedge between future promised utilities conditional on employment and unemployment is required to keep the IC constraint. In contrast, when monitoring is used, the planner uses the information of both the employment and the signal outcomes. This allows her to reduce the

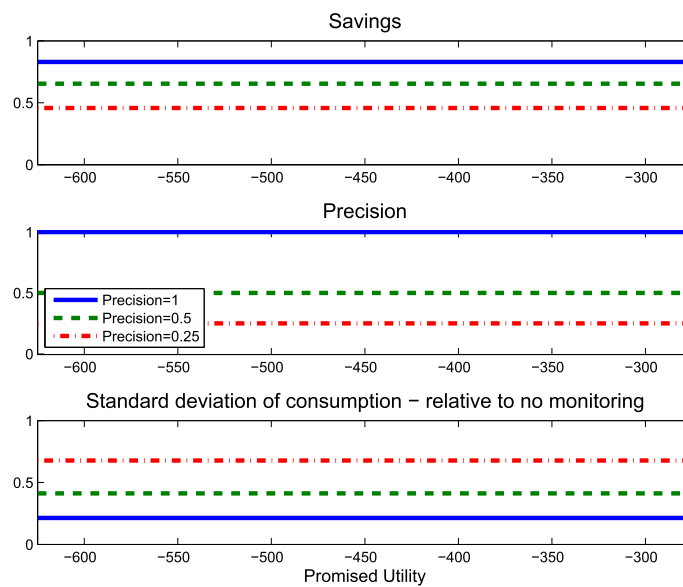


FIGURE 4. Statistics for two effort levels and exogenous precision.

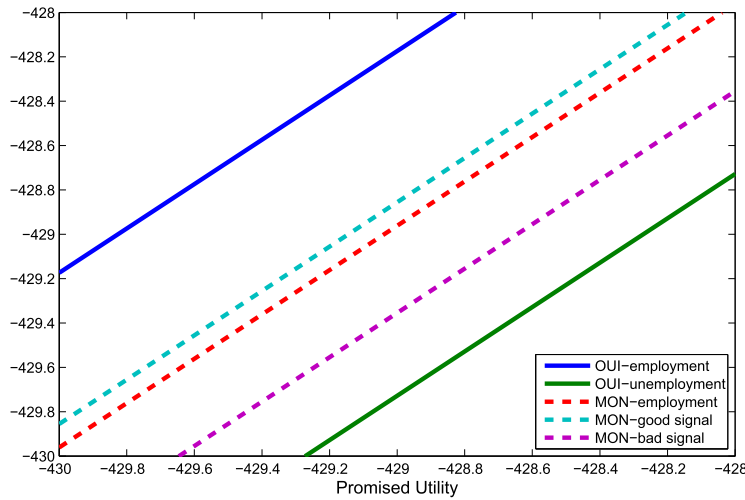


FIGURE 5. Spreads for two effort levels and exogenous precision.

spread between promised utilities, which in turn translates into smaller consumption volatility over time for each worker.

Another way to look at this result regarding the spread is the following. Consumption variation stems from both variation within a state (e.g., within unemployment) and across states (between employment and unemployment). Monitoring increases consumption variation within unemployment. This may seem surprising given that the worker is risk averse. But, since this is done based on additional information, by doing so the planner is able to reduce consumption variation across states. Indeed, Figure 5 shows clearly that (a) consumption variation within unemployment increases (from no variation when there is no monitoring to variation that stems from the response to the signal) and (b) consumption variation across states (employment and unemployment) is much reduced.

A convenient and informative metric of the volatility of consumption is the standard deviation of consumption paths that are simulated using the optimal monitoring and UI policies. I simulate at several levels of promised utility 1000 workers over 60 periods using the exogenous separation shocks and the job-finding probability that is associated with the optimal effort level. I then determine the optimal consumption and continuation promised utility for each worker at each period. Finally, I calculate the average standard deviation of those consumption paths for workers under both optimal UI and monitoring. To show the relative importance of monitoring, I depict the ratio of the standard deviation of consumption under monitoring to that under optimal UI. Since monitoring can only decrease the standard deviation of consumption, which is the sole purpose of using monitoring, that ratio is bounded above by 1. This, together with the

fact that optimal UI is associated with a positive standard deviation, implies that the ratio of standard deviations is bounded by zero from below.²⁵

The bottom panel of Figure 4 shows the ratio of standard deviations of consumption for three different levels of precision. As the precision increases, the ability of the planner to reduce the volatility of consumption increases, resulting in a lower consumption volatility.

The main takeaway from this simple case is that an increased precision leads to higher savings through a lowered variance of consumption.

5.4.2 Two effort levels with endogenous monitoring The previous case with exogenous precision was used to illustrate the role of information for a model with additional information on the worker’s action. Taking the analysis further, I now assume that monitoring is costly and that the precision at each state is chosen optimally by the planner. This has important implications for the results. In Figure 6, which summarizes the results for this case, I zoom in on a partial interval of the variables’ support in the vertical axes in the top and bottom panels to emphasize the variables’ profiles over promised utility.

Consider first the savings fraction in the top panel of Figure 6. As expected, relative to the case in Figure 4, in which precision is equal to one, the average savings here are lower because now the planner incurs the cost of the additional information. Notice that the savings associated with monitoring increase with promised utility, even over intervals

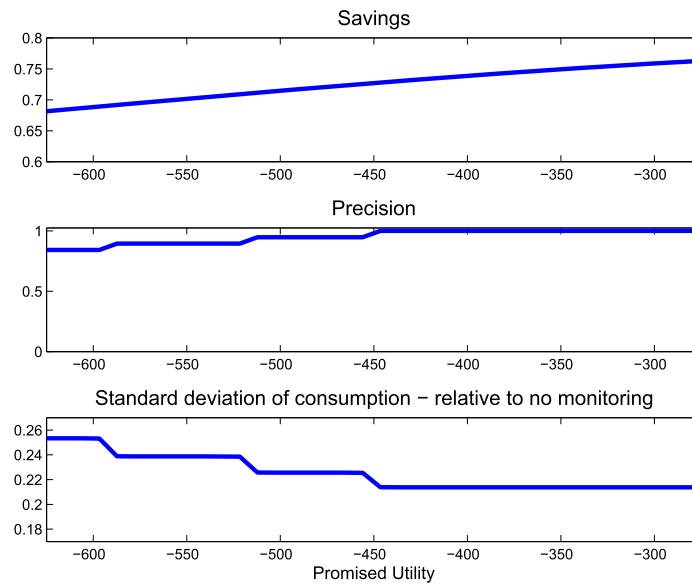


FIGURE 6. Statistics for two effort levels and endogenous precision.

²⁵Notice that the results of monitoring can be then characterized by three fractions: the savings associated with monitoring, the precision (whether exogenous or endogenous), and the standard deviation of consumption under monitoring relative to that of optimal UI.

where the optimal precision (middle panel) is constant.²⁶ As we saw before, the cost of moral hazard increases with promised utility. Since the cost of precision is independent of promised utility, even when precision is fixed, the cost of precision relative to the cost of moral hazard decreases and, therefore, the savings increase with promised utility.

Importantly, and related to the previous point, the fact that the cost of moral hazard (and thus the potential for savings) increases with promised utility calls for choosing a higher precision level for higher levels of promised utility. This can be seen in the middle panel of Figure 6, where the optimal precision's profile is mildly increasing over promised utility. The bottom panel of Figure 6 shows the volatility of consumption relative to promised utility. As the precision increases, the volatility of consumption decreases.

5.4.3 A continuous effort with exogenous monitoring A continuous effort adds one more decision variable to the planner's choice set: the optimal effort level.²⁷ The top panel of Figure 7 shows the optimal effort level under the first best, optimal UI, and monitoring over promised utility. Two clear properties of the solutions emerge from

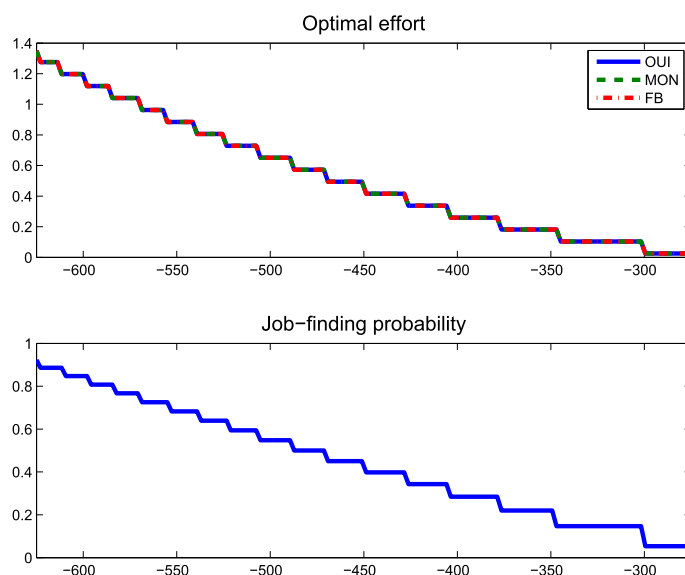


FIGURE 7. Optimal effort for a continuous effort and exogenous precision.

²⁶The intervals with constant precision are a result of the discrete representation of effort in the quantitative analysis. Increasing the number of grid points for the effort in the numerical solution would make the profile smoother. However, the constant levels of precision are interesting as they show that (1) savings increase with promised utility even when (costly) precision is fixed, and (2) standard deviation of consumption is constant when precision is constant.

²⁷As described above in the model with two effort levels, it is never optimal to monitor a worker who became employed, as the employment outcome reveals with certainty that he exerted the recommended effort. When effort is continuous, this is no longer the case. Nevertheless, I abstract from monitoring employed workers as it substantially complicates the analysis without providing additional insights. Moreover, notice that monitoring of employed workers is not practiced in the United States or elsewhere.

that figure. First, according to the utility form $u(c) - a$, the higher the promised utility, the higher the compensation required for compensating the worker for her effort. This *wealth* effect leads to effort dynamics that span almost the entire job-finding probability (bottom panel).

Second, the optimal effort level is indistinguishable among the three problems: the first best, optimal UI, and monitoring. Although we expect that the additional information available to the planner would allow supporting a higher effort level, this effect is dwarfed by the wealth effect. This result is consistent with the observation of [Hopenhayn and Nicolini \(1997\)](#) that the allocation of optimal unemployment insurance gets quite close to that of the first best.

Figure 8 presents the savings (top panel) for several exogenous levels of precision (middle panel). Comparing the three curves associated with the three precision levels shows that, as in the previous case, the higher the quality of the information associated with monitoring, the higher the gain from monitoring. Similarly, maximum precision does not imply the first best's value. The reason is that even though the information in this scenario is costless, it is not perfect.

The profile of savings, however, is changed with respect to the previous scenarios, presenting an inverse U-shaped curve instead of a monotone increasing profile. An important implication of the wealth effect discussed above is that as optimal effort decreases, so does the value of employment, because the worker's compensation for her job-search effort increases with the promised utility. For a high enough promised utility, employment becomes undesired, even under the first best, because the consumption

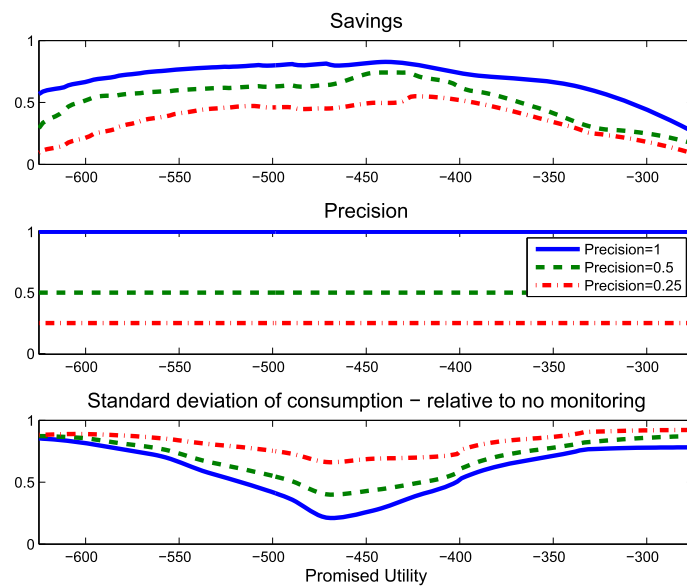


FIGURE 8. Statistics for a continuous effort and exogenous precision.

required to compensate the worker for her work effort is higher than the wage.²⁸ In the limit, where employment, even under the first best, is undesired, precision should be zero. This effect implies that precision should *decrease* with promised utility. Notice that this driving force was not present in the analysis with only two effort levels.

In contrast, the increase in the cost of moral hazard with promised utility, which was already present in the analysis with two effort levels, implies that precision should *increase* with promised utility. The combination of those two economic forces delivers an inverse U-shaped profile for savings: for low levels of utility, the cost of moral hazard is low and the value of information is low; for intermediate values of promised utility, where the cost of moral hazard is high and employment has a substantial value for the planner, the savings are high; finally, for high levels of promised utility, the wealth effect makes employment undesired regardless of the existence of moral hazard, and again the gain is low.

The standard deviation ratio shown in the bottom panel of Figure 8 shows another change in pattern from previous scenarios, exhibiting a U-shaped pattern with respect to promised utility. Once again, the difference here is due to the choice of job-search effort. In the previous scenarios, the effort levels of employment and unemployment were the same. Here, in contrast, the large differences in effort call for large differences in consumption. Specifically, for low levels of promised utility, where the job-search effort is higher than the work effort, consumption goes up each time the worker becomes unemployed to compensate her for the increase in effort. For high levels of promised utility, the opposite happens: unemployment is associated with a lower effort, so lower consumption during that stage is needed to keep the utility promised and to incentivize the worker to search for a job. Therefore, in the two extremes, a signal has a limited ability to reduce consumption variation.

5.4.4 A continuous effort with endogenous monitoring Armed with the insights from the previous simplified analyses, we are now ready for the realistic case of monitoring with a continuous effort and endogenous (costly) monitoring. Since the optimal effort is unaffected by the endogeneity of the precision and is identical to that shown in Figure 7, I proceed to describe the planner's other choices.

Figure 9 shows the main result of the full case. The middle panel shows the optimal precision level. The forces that shape the importance of information discussed above lead to the inverse U-shaped pattern. On the one hand, the decrease in the value of employment with promised utility implies that precision should *decrease* with U . On the other hand, the fact that the importance of moral hazard increases with promised utility, implies that precision should *increase* with U . Notice that for high enough and low enough values of promised utility, the optimal precision is zero. The top panel presents the savings associated with optimal precision, reflecting the same economic forces that shape the optimal precision. Quantitatively, savings go up to 50%, but as the analysis

²⁸In the restrictive two-effort model, the planner cannot reduce the worker's effort to find employment in order to account for the fact that the value of employment is decreasing. Notice that for a high enough promised utility, supporting incentives to search (via the single probability) is not beneficial anymore. However, for the setup used in the quantitative analysis, this is not the case.

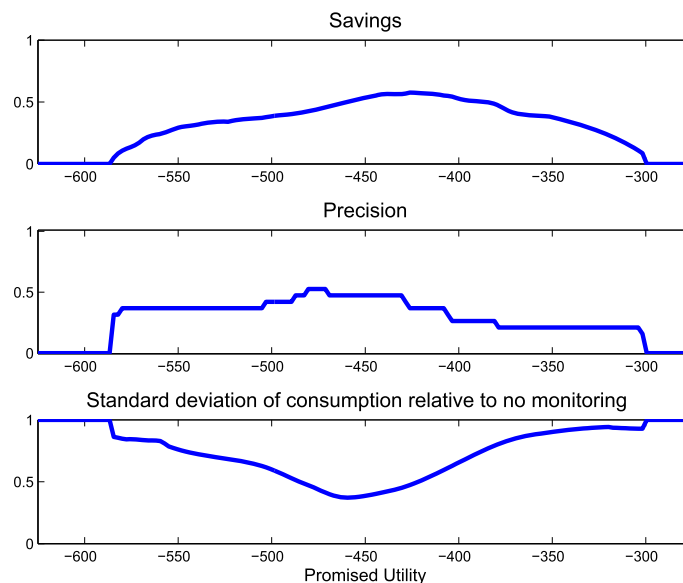


FIGURE 9. Statistics for a continuous effort and endogenous precision.

above has shown, the importance of monitoring is strongly dependent on the generosity of the welfare system.

The bottom panel of Figure 9 shows the standard deviation of consumption under monitoring relative to optimal UI. Although the precision is not constant as it is in Figure 8, the profile of the standard deviation is similar, reflecting the relative importance of the difference between work and job-search effort levels.

5.4.5 Consumption over the unemployment spell A key property in the optimal UI literature is that consumption over the unemployment spell should monotonically decrease (see, e.g., [Hopenhayn and Nicolini \(1997\)](#)). The reason for that pattern is simple. Each period of unemployment results in a lower promised utility for the worker. Since there is a monotonic relationship between promised utility and consumption, as long as the worker is unemployed, she will experience a drop in her consumption. That drop in consumption is required at each period to maintain her incentive to search for a job with the effort recommended by the planner.

At the other extreme, the first-best allocation allows the planner to provide the worker with smooth consumption over states and time. What does the consumption of the model with monitoring look like? We already know from the analysis above that monitoring reduces the standard deviation of consumption relative to optimal UI. To illustrate the concrete implications of that reduction, I simulate the model for several scenarios of monitoring and study how the dynamics of consumption is affected by monitoring over time for each scenario.

Figure 10 shows the optimal consumption level for three scenarios of seven periods each for the model with continuous effort and endogenous monitoring. In the first four periods of each scenario, the worker is unemployed; that is, she does not find a job even

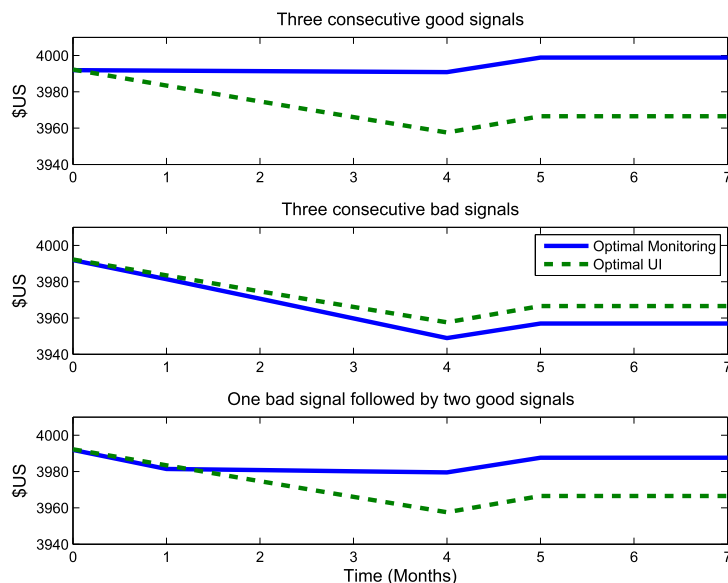


FIGURE 10. Conditional simulations.

though she exerts the recommended optimal effort. At the fifth period, she finds a job and remains employed through the seventh period. Since in UI (without monitoring), there is only one outcome of unemployment, consumption under that policy (shown by the dashed line) is identical in the three scenarios. In contrast, under the monitoring policy there are two possible unemployment outcomes, depending on whether the monitoring signal produced was good or bad. In the top panel of Figure 10, the worker experiences three consecutive good signals. Her consumption is almost unchanged before increasing in response to employment. In the second scenario, she experiences three consecutive bad signals, leading to a reduction in consumption that is a bit stronger than that of UI. In the bottom panel, where the worker first receives a bad signal and then receives two good signals, consumption first drops, then is almost unchanged and then increases upon employment.

There are several lessons to take from this exercise. First, given that a typical worker receives both good and bad signals, monitoring mitigates the consumption volatility associated with UI. Second, the extent to which consumption changes in response to a good signal depends on the parameters of the model and, especially, those of the signal. In the numerical example shown in the figure, consumption upon a good signal very slightly decreases. In other experiments, for which figures are not provided, increasing θ from 0.6 to 0.9 leads to consumption *increasing* upon a good signal. Therefore, while consumption monotonically decreases under UI, it does not necessarily decrease when monitoring is present.²⁹

²⁹Notice the difference between the optimal contract approach and the one used in constrained efficiency models such as the one by Boone et al. (2007). In optimal contracts, the level of consumption continuously updates based on all the relevant information that is available for the planner. In the alternative

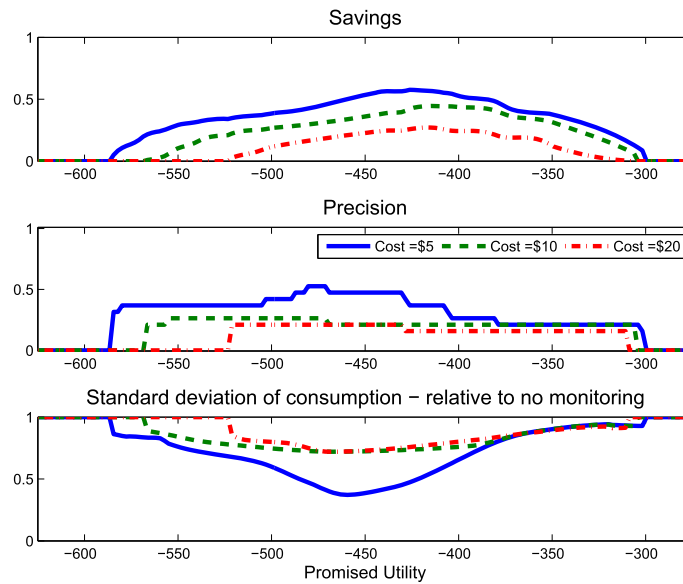


FIGURE 11. Variation in monitoring cost.

5.4.6 *Sensitivity analysis with respect to monitoring cost* I conclude this subsection with a sensitivity analysis of the model's results to the cost of monitoring. In the analysis above, I used a cost of monitoring at the calibrated signal of \$5. Figure 11 shows the results for two additional costs: \$10 and \$20. It is evident from the middle panel that when the monitoring cost increases the planner uses a less precise signal. The bottom panel shows the standard deviation of consumption from simulations using the optimal policy for each cost: as we saw above, the signal is valuable to the planner because it allows her to reduce the consumption variation. As the cost increases and the signal chosen is less precise, the standard deviation of consumption under monitoring gets closer to that of optimal UI, where monitoring is unavailable. Finally, the top panel shows that savings declines with the cost of monitoring: for low promised utilities, the savings decreases because as the cost increases and the signal's precision decreases, the planner needs to compensate the worker more for the consumption variation; for high promised utilities even though the signal does not change much, savings decreases simply due to the higher cost of monitoring.

6. FAKING SEARCH BEHAVIOR

The model studied so far assumed that the worker's sole decision is her job-search effort. In reality, the worker may also affect the monitoring signal by faking her search behavior. In this section, I extend the one-period model to quantitatively study this variant of the model.

approach, consumption takes a finite number of values that are updated based only on the current state of the worker and the current signal on her job search.

Assume that the worker can exert an effort $f(\eta)$ in order to receive a good signal with probability $\eta \in [0, 1]$, where $f(\eta) = \zeta \eta^\phi$, with $\phi > 1$, and $f(0) = 0$.

The planner's problem in the one-period setting now becomes:

$$\begin{aligned}
 V(U) &= \max_{c^e, c^g, c^b, \theta, \eta} \{ \pi(w - c^e) - (1 - \pi)\theta c^g - (1 - \pi)(1 - \theta)c^b - (1 - \pi)\kappa(\theta) \} \\
 \text{s.t. :} \\
 U &= \pi u(c^e) + (1 - \pi)\theta u(c^g) + (1 - \pi)(1 - \theta)u(c^b) - e \\
 &\quad - f(\eta) + \eta u(c^g) + (1 - \eta)u(c^b) \leq \pi u(c^e) + (1 - \pi)\theta u(c^g) \\
 &\quad + (1 - \pi)(1 - \theta)u(c^b) - e, \\
 \eta &\in \arg \max_{\eta} \{ -f(\eta) + \eta u(c^g) + (1 - \eta)u(c^b) \}, \tag{14}
 \end{aligned}$$

where the left-hand side of the second constraint (the IC) takes into account the probability of a good signal given the faking behavior, and the additional constraint is the worker's optimal choice of η . Notice that the IC constraint to induce search becomes redundant as zero effort is a special case of the IC constraint in Problem 14, with $\eta = 0$.

I use this extended model for two purposes. First, I show quantitatively that the main result of Proposition 1 holds for the commonly used preferences class of CRRA. Second, I investigate how the presence of faking search affects the optimal decisions of the planner and how the faking behavior itself changes with the worker's state of promised utility U .

Proposition 1 states that the monitoring precision increases and the utility spreads decrease if and only if the derivative of the inverse of the worker's utility is convex. As a starting point for the analysis, I use the knife-edge case of CRRA with a coefficient of constant relative risk aversion of 0.5, where the derivative of the inverse is linear.

Figure 12 shows the planner's and the worker's optimal decisions for CRRA with $\sigma = 0.5$ with and without faking search effort.³⁰ Let us first analyze the case without search effort (in solid lines). In this case, the probability of a good signal given no search effort (η) is by construction zero (bottom panel), and the signal's precision (θ) is around 0.5. The second panel shows the standard deviation of the utility delivered to the worker conditional on the outcome. In this case, which was analyzed in Section 4, the monitoring precision and the spread are independent of the utility promised to the worker.

Turning to the case where the worker can fake her search effort (in dashed lines), we first observe that the monitoring precision, the spreads, and the faking behavior are all independent of the utility promised to the worker exactly as predicted by Proposition 1 for the model without faking search effort. The question then is how the faking search behavior affects the planner's decision on applying the monitoring technology.

³⁰Using the parameters of the infinite horizon in this section results, in some of the cases, in a corner solution for the monitoring precision. To best illustrate the mechanisms at work, I choose parameters that deliver an internal solution. See Appendix C.1 for details. The qualitative results are insensitive to the parameters (except, of course, for the corner solutions).

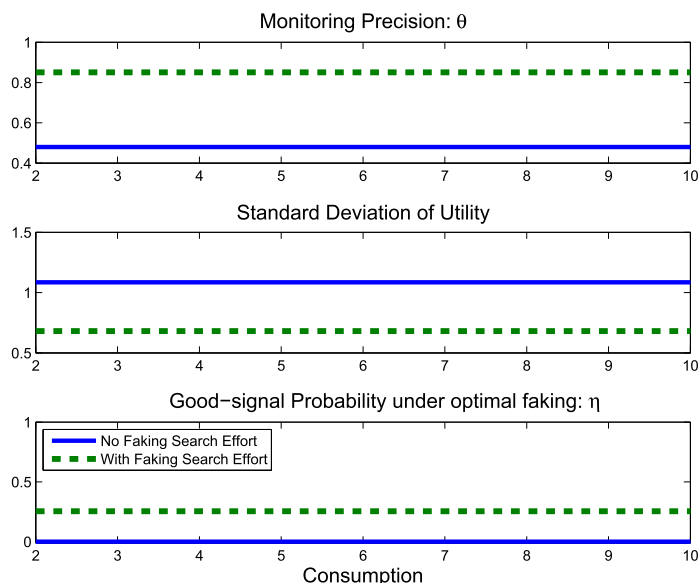


FIGURE 12. Optimal choices with and without faking search effort for $\sigma = 0.5$.

When faking search effort is possible, the standard deviation of utilities motivates the worker to fake her search behavior. This means that the planner faces an additional cost for spreading out utilities, as it encourages the worker to fake her search behavior. This leads the worker to reduce the spread in utility (middle panel) and instead increase the monitoring precision in order to preserve the incentive-compatibility constraint of the worker to search.

The take away from this case is that given the worker’s search behavior the planner shifts away from using a spread to increasing the precision of monitoring.

We can now move away from the knife-edge case of CRRA with $\sigma = 0.50$ to other preferences. I study two additional cases of CRRA, with coefficients of relative risk aversion, σ , that are equal to 0.75 (where the derivative of the inverse of the worker’s utility is strictly convex), and to 0.25 (where the derivative of the inverse of the worker’s utility is strictly concave).

Figure 13 shows the planner’s and the worker’s optimal decisions for CRRA with $\sigma = 0.75$ with and without faking search effort. First, observe that the result achieved in the previous case (with $\sigma = 0.5$) is maintained for this case as well: the faking behavior of the worker leads the planner to shift away from the spread (middle panel) and to increase the monitoring precision (top panel). Also note that the driving force behind Proposition 1, that the cost of spreading out utility increases with utility for CRRA with a coefficient greater than 0.5, is still present where faking is possible. This explains the increase in precision in the top panel and the decrease in the spread in the middle panel. Finally, the increase in the spread over utility decreases the incentive of the worker to fake her search behavior and as a consequence η decreases over utility (bottom panel).

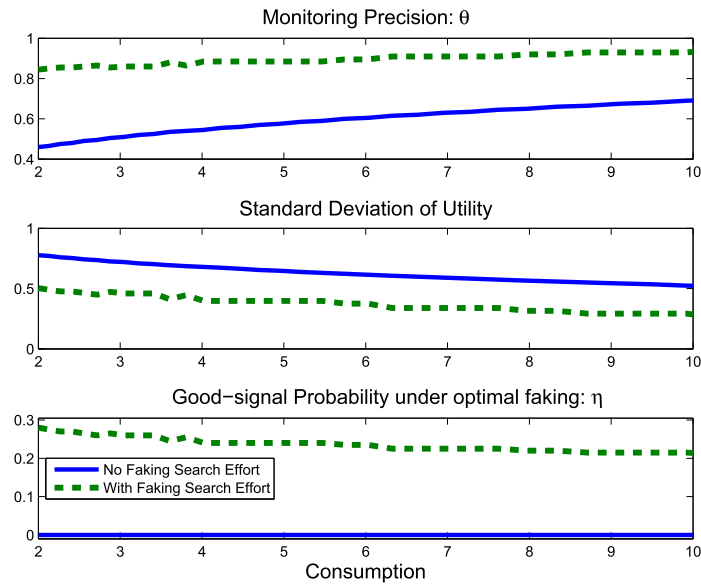
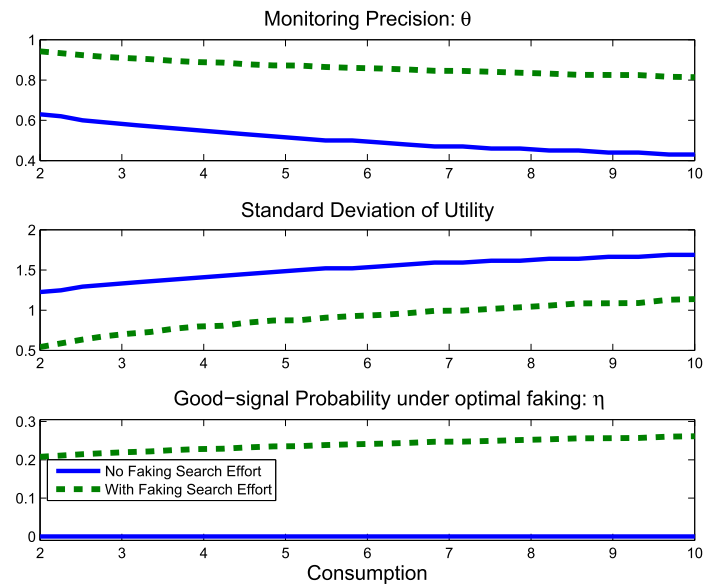
FIGURE 13. Optimal choices with and without faking search effort for $\sigma = 0.75$.FIGURE 14. Optimal choices with and without faking search effort for $\sigma = 0.25$.

Figure 14 shows the planner's and the worker's optimal decisions for CRRA with $\sigma = 0.25$ with and without faking search effort. The two insights from above hold for this case as well. First, with faking behavior the planner decreases the spread and increases the monitoring precision. Second, the main result of Proposition 1 holds as the planner

decreases the monitoring precision and increases the spread over utility for this case where the inverse of the worker's utility has a *concave* derivative.

Figures 15 and 16 in Appendix C.1 show that the same results are qualitatively maintained when going closer to the knife-edge case for CRRA with relative risk aversion of $\sigma = 0.55$ (convex first derivative) and $\sigma = 0.45$ (concave first derivative).

7. CONCLUDING REMARKS

In this paper, I add job-search monitoring to the optimal unemployment insurance framework. The introduction of monitoring into the model follows the practice of monitoring in the United States: a caseworker verifies the job-search activity of an unemployed worker (in the form of employment contacts) with some precision and sanctions the worker if the effort seems unsatisfactory.

Allowing the signal to be imperfect in this analysis has important advantages both qualitatively and quantitatively. Qualitatively, the optimal contract includes three realistic features: a nontrivial decision of the monitoring precision; endogenously limited sanctions and rewards; and application of sanctions in equilibrium. Quantitatively, this technology permits an assessment of the optimal monitoring technology and its value.

I use the standard assumption in optimal unemployment insurance that the hazard rate is constant over the unemployment spell. Under this assumption, which allows a clear characterization of the contract, the monitoring precision and the sanction are fairly constant along the unemployment spell. In future research, the model can be extended to allow for negative duration dependence (as in Pavoni (2009)) in order to study the dynamics of monitoring frequency and sanctions over the unemployment spell.

One limitation of the framework used in this paper is that the model's tractability depends on the assumption that the planner controls the worker's consumption, that is, that no savings are allowed on the worker's side. As pointed out by Abdulkadiroglu, Kuruscu, and Şahin (2002) and Shimer and Werning (2008), allowing workers to accumulate unobservable savings may significantly affect the results. The recursive contract framework strength is in demonstrating the main trade-offs for the optimal contract when a costly imperfect signal is available. It appears that as long as differentiation of future payments is necessary, monitoring could be effective in reducing the need for costly spreads.

Another limitation of this framework is that the sanctions are unjustified. This occurs because the IC constraint in the model holds. Nevertheless, the sanctions need to be placed in the contract to keep the worker's incentives in place. Note that the same concept of unjustified punishment holds in optimal UI. There, conditional on unemployment, the worker experiences benefit cuts even though the planner is aware that the worker put forth the recommended effort. A more realistic model would include unobserved heterogeneity in disutility from job search and from work. Under such circumstances, the sanctions in equilibrium would be partially justified.

In reality, workers are not only choosing effort but also type of job that they are applying to. The optimal unemployment insurance literature defines a contract between the government and the worker that is focused on the job-search effort. Hence, the trade off is centered at the moral hazard problem of hidden information. The effect of including a choice of a job in the model depends on how this additional choice is modeled. As

long as there is consensus between the planner and the worker as to what type of job the worker should be looking for, all the results of that literature go through. This can be done by having the planner choosing for the worker the type of job that she should be looking for, taking into account the optimal contract associated with each type of job. If this is not the case, and additional concerns are to be addressed, then this would require much attention on the modeling part. The reason is that a key assumption in the optimal UI literature is that the planner takes over the wage of the worker. Therefore, the contract does not and cannot condition on the wage but only on the employment outcome.

According to Grubb (2000), there are significant differences across countries in all the policies' main characteristics. For example, in Australia, a moderate sanction of 18% of the benefits level is applied for a duration of 6 months, which is considerably longer than the one-week denial of benefits in the United States. At the same time, Australia's annual sanction rate, standing at 1.2%, is relatively low when compared with the 33% in the United States. An extended model could reveal whether the variation in policies follows labor market characteristics or some inefficiencies.

APPENDIX A: PROOFS

CLAIM 1. *In the optimal solution, $u(c^e) = u(c^g) > u(c^b) = U$.*

PROOF. The proof is a combination of Lemmata 1, 2, and 4 below. \square

LEMMA 1. *In the optimal solution, either $c^e > c^b$ or $c^g > c^b$, or both.*

PROOF. Rewrite the IC as

$$\pi u(c^e) + (1 - \pi)\theta u(c^g) \geq [\pi + (1 - \pi)\theta]u(c^b) + e. \quad (15)$$

Since $e > 0$ and since the sum of the coefficients of $\{u(c^e), u(c^g)\}$ is equal to the coefficient of $u(c^b)$ (and positive), if both $c^e \leq c^b$ and $c^g \leq c^b$ then the IC cannot hold. \square

LEMMA 2. *In the optimal solution, $c^e = c^g$.*

PROOF. Assume that $c^e > c^g$. The optimal solution can be improved as follows. Decrease c^e by ε and increase c^g by $\frac{\pi\varepsilon}{(1-\pi)\theta}$.³¹ By construction, this change does not affect the objective function because $\pi(c^e - \varepsilon) + (1 - \pi)\theta(c^g + \frac{\pi\varepsilon}{(1-\pi)\theta}) = \pi c^e + (1 - \pi)\theta c^g$. To study the effect on the PK, consider the part of the PK composed of $\pi u(c^e) + (1 - \pi)\theta u(c^g)$. The change increases the right-hand side of the PK because it is a lottery with the same certainty equivalent but with less risk. Formally, the claim is that:

$$\begin{aligned} \pi u(c^e - \varepsilon) + (1 - \pi)\theta u\left(c^g + \frac{\varepsilon\pi}{(1-\pi)\theta}\right) &> \pi u(c^e) + (1 - \pi)\theta u(c^g), \\ (1 - \pi)\theta\left(u\left(c^g + \frac{\varepsilon\pi}{(1-\pi)\theta}\right) - u(c^g)\right) &> \pi(u(c^e) - u(c^e - \varepsilon)). \end{aligned} \quad (16)$$

³¹This line of proof works only if $\theta > 0$ as I divide by θ . If $\theta = 0$, then c^g does not affect the planner's value or the constraints and it can be arbitrarily assumed to be equal to c^e . Notice that $\pi \in (0, 1)$ by assumption.

Divide both sides by $\varepsilon > 0$ and rearrange to get

$$\frac{u\left(c^g + \frac{\varepsilon\pi}{(1-\pi)\theta}\right) - u(c^g)}{\frac{\varepsilon\pi}{(1-\pi)\theta}} > \frac{u(c^e) - u(c^e - \varepsilon)}{\varepsilon}. \quad (17)$$

In the limit, this is $u'(c^g) > u'(c^e)$, which is true by the negation assumption that $c^g < c^e$. Thus the right-hand side of the PK becomes larger than U . As the right-hand side of the IC is the same as that of the PK, the right-hand side of the IC increases as well, making the IC slack (or slacker). For the same reason, the IC becomes slack (or slacker); see (15).

Decrease c^g by $\delta < \frac{\pi\varepsilon}{(1-\pi)\theta}$ such that the PK is tight again. The IC still holds. This additional deviation increases the objective function, a contradiction to the solution being optimal.

Analogous arguments work for the case of $c^g > c^e$. \square

LEMMA 3. *In the optimal solution, the IC holds with equality.*

PROOF. By Lemmata 1 and 2 $c^e > c^b$. If the IC is slack, then the objective function can be improved. Decrease c^e by ε and increase c^b by $\frac{\varepsilon\pi}{(1-\pi)(1-\theta)}$. (Notice that $\theta < 1$ since $\lim_{\theta \rightarrow 1} \kappa(\theta) = \infty$.) By construction, this change does not affect the objective function. The IC still holds. Consider the part of the PK composed of $\pi u(c^e) + (1-\pi)(1-\theta)u(c^b)$. Those changes make the PK slack because it is a lottery with the same certainty equivalent but with less risk. Formally, the claim is that:

$$\begin{aligned} & \pi u(c^e - \varepsilon) + (1-\pi)(1-\theta)u\left(c^b + \frac{\varepsilon\pi}{(1-\pi)(1-\theta)}\right) \\ & > \pi u(c^e) + (1-\pi)(1-\theta)u(c^b), \\ & (1-\pi)(1-\theta)\left(u\left(c^b + \frac{\varepsilon\pi}{(1-\pi)(1-\theta)}\right) - u(c^b)\right) > \pi(u(c^e) - u(c^e - \varepsilon)). \end{aligned} \quad (18)$$

Divide both sides by ε and rearrange to get

$$\frac{u\left(c^b + \frac{\varepsilon\pi}{(1-\pi)(1-\theta)}\right) - u(c^b)}{\frac{\varepsilon\pi}{(1-\pi)(1-\theta)}} > \frac{u(c^e) - u(c^e - \varepsilon)}{\varepsilon}. \quad (19)$$

For a sufficiently small epsilon, this is true since $u'(c^b) > u'(c^e)$ (recall that $c^b < c^e$). Now decrease c^e by $\delta < \varepsilon$ in order to improve the objective function without damaging any of the constraints. \square

LEMMA 4. *In the optimal solution, $u(c^b) = U$.*

PROOF. Since both the PK and the IC are tight, and since the left-hand side of both constraints is identical, the right-hand side of both constraints is equal and $u(c^b) = U$. \square

PROPOSITION 1. *The solution to Problem 2 has the following characteristics if $(u^{-1})'(\cdot)$ is convex:*

- (i) *the optimal signal's precision (θ) increases with promised utility (U);*
- (ii) *the utility spread ($\frac{e}{\pi+(1-\pi)\theta}$) decreases with promised utility (U);*
- (iii) *the cost of spreading out utility increases with the promised utility (U);*
- (iv) *the converse version of (i)–(iii) holds when $(u^{-1})'(\cdot)$ is concave.*

PROOF. The three parts of the proposition are proved sequentially.

Proof of (i)

The proof is based on monotone comparative statics (Milgrom and Shannon (1994)):

$$\begin{aligned}\frac{\partial V}{\partial U} &= \alpha \left((u^{-1})' \left(U + \frac{e}{\alpha} \right) - (u^{-1})'(U) \right) + (u^{-1})'(U), \\ \frac{\partial^2 V}{\partial U \partial \theta} &= (1 - \pi) \left((u^{-1})' \left(U + \frac{e}{\alpha} \right) - (u^{-1})'(U) \right) - \frac{e(1 - \pi)}{\alpha} (u^{-1})'' \left(U + \frac{e}{\alpha} \right).\end{aligned}\tag{20}$$

According to the monotone comparative statics theorem, θ^* (weakly) increases with U if $\frac{\partial^2 V}{\partial U \partial \theta} \leq 0$. This is satisfied if $(u^{-1})'(\cdot)$ is convex. To see this, rewrite $\frac{\partial^2 V}{\partial U \partial \theta}$ as

$$\frac{e(1 - \pi)}{\alpha} \left\{ \frac{(u^{-1})' \left(U + \frac{e}{\alpha} \right) - (u^{-1})'(U)}{\frac{e}{\alpha}} - (u^{-1})'' \left(U + \frac{e}{\alpha} \right) \right\},\tag{21}$$

and notice that $\frac{(u^{-1})' \left(U + \frac{e}{\alpha} \right) - (u^{-1})'(U)}{\frac{e}{\alpha}}$ is the slope of $(u^{-1})'(\cdot)$ between $\{U, U + \frac{e}{\alpha}\}$, where $\frac{e}{\alpha} > 0$, and $(u^{-1})'' \left(U + \frac{e}{\alpha} \right)$ is the slope of $(u^{-1})'(\cdot)$ at $(U + \frac{e}{\alpha})$. If $(u^{-1})'(\cdot)$ is convex, then the slope of $(u^{-1})'(\cdot)$ at $(U + \frac{e}{\alpha})$ is higher than the slope between $\{U, U + \frac{e}{\alpha}\} \Rightarrow$ (21) is negative $\Rightarrow \frac{\partial^2 V}{\partial U \partial \theta} \leq 0$.

Proof of (ii)

By (i), optimal θ increases with U if $(u^{-1})'(\cdot)$ is convex. The spread is $\frac{e}{\pi+(1-\pi)\theta}$, so the spread decreases as θ increases.

Proof of (iii)

The planner's cost in Problem 2 is composed out of the cost of providing a consumption, equal to $\alpha u^{-1} \left(U + \frac{e}{\alpha} \right) + (1 - \alpha) u^{-1}(U)$ and the monitoring cost $(1 - \pi) \kappa(\theta)$.

The first best cost for the planner is $u^{-1}(U + e)$. Therefore, the difference between the planner's cost of providing a consumption in the first best and in (2) is a result of the requirement of spreading out utilities. Define $D(U)$ as the cost of spreading out utilities $\{U, U + \frac{e}{\alpha}\}$ at utility U as the difference between the costs

$$D(U) = \left\{ \alpha u^{-1} \left(U + \frac{e}{\alpha} \right) + (1 - \alpha) u^{-1}(U) \right\} - u^{-1}(U + e),\tag{22}$$

and notice that the curly brackets include a lottery with prizes $\{U + \frac{e}{\alpha}, U\}$ with probabilities $\{\alpha, 1 - \alpha\}$, whose expected prize is $U + e$. This means that $D(U)$ is the difference

between a lottery and a certainty equivalent $(U + e)$, valued by the function of $u^{-1}(\cdot)$. Since $u(\cdot)$ is concave $D(U) > 0 \forall u(\cdot)$.

The dependence of this cost on U is the following derivative:

$$D'(U) = \left\{ \alpha(u^{-1})' \left(U + \frac{e}{\alpha} \right) + (1 - \alpha)(u^{-1})'(U) \right\} - (u^{-1})'(U + e). \quad (23)$$

Since under Condition 1 $(u^{-1})'$ is convex, Jensen's inequality implies that

$$\begin{aligned} \alpha(u^{-1})' \left(U + \frac{e}{\alpha} \right) + (1 - \alpha)(u^{-1})'(U) &> (u^{-1})' \left\{ \alpha \left(U + \frac{e}{\alpha} \right) + (1 - \alpha)U \right\} \\ &= (u^{-1})'(U + e) \\ &\Rightarrow D'(U) > 0. \end{aligned}$$

Proof of (iv)

The proof follows the same arguments as in the proof for parts (i)–(iii) above for $(u^{-1})'(\cdot)$ concave. \square

APPENDIX B: COMPUTATIONAL METHOD

To lighten notation, this section is written for a general signal structure. Hence, $\min(1, \max(0, \underline{\theta} + \zeta \cdot (e - \underline{e}))$ is replaced by $\theta(\underline{e}, e)$. $\underline{\theta}$ still represents the probability of a good signal given that the recommended effort is actually exerted.

We begin by transforming Problem 4, which has six decision variables and two constraints, into a problem with four variables (U^g , U^b , θ , and e) and no constraints. We then enumerate on two variables (θ and e) and solve numerically for the remaining two variables (U^g , U^b). Unlike the support for the continuation values, which is the real line, the support for precision and effort are more restricted. Numerically, it is more efficient to enumerate over dense values of those variables rather than including constraints on their state space.³²

We continue by using the IC, which is

$$e \in \operatorname{argmax}_{e \in [0, \infty)} \{ u(c) - e + \beta \pi(e) U^e + (1 - \pi(e)) [\theta(\underline{e}, e) U^g + (1 - \theta(\underline{e}, e)) U^b] \},$$

³²Notice that since the continuation values lie on the real line, the dynamic programming results do not hold. This is a known complication in the optimal unemployment insurance literature, such as in [Hopenhayn and Nicolini \(2009\)](#). I handle this by imposing that the continuation utilities lie in a compact set. This solves the issue but admittedly solves a slightly different model than the one described. To get more confidence that this limitation is not crucial for the quantitative results, I relax this assumption and extrapolate over promised utilities to confirm that indeed the results are unchanged.

to derive the first-order condition of the worker's problem as³³

$$1 = \beta \pi'(e)U^e + (1 - \pi'(e))[\theta(\underline{e}, e)U^g + (1 - \theta(\underline{e}, e))U^b] \\ + (1 - \pi(e))\theta'(\underline{e}, e)[U^g - U^b]. \quad (24)$$

From this condition, U^e can be expressed as a function of U^g and U^b , as follows:³⁴

$$U^e = \frac{1}{\beta \pi'(e)} \{1 - (1 - \pi'(e))[\theta(\underline{e}, e)U^g + (1 - \theta(\underline{e}, e))U^b] \\ - (1 - \pi(e))\theta'(\underline{e}, e)[U^g - U^b]\}. \quad (25)$$

Using the PK, c can be stated as a function of U^g , U^b , and U^e , as

$$U = u(c) - e + \beta \pi(e)U^e + (1 - \pi(e))[\underline{\theta}U^g + (1 - \underline{\theta})U^b]. \quad (26)$$

Since U^e is given by equation (25), c can be written as a function of only U^g and U^b .

The following pseudo algorithm describes the convergence of V given an existing guess for W . Once V converged for the current solution of W , iterate again on W till its convergence given the updated guess for V and so on until a certain threshold on both V and W across iterations is reached.³⁵

- (i) Construct a grid of V over promised utility U , signal's precision θ , and effort e .
- (ii) Use the last guess for V , or an initial guess if one is not available. A simple and numerically effective guess is the first best using work effort.
- (iii) Enumerate the following over the cross grid of effort and precision:
 - (a) Solve numerically (and jointly) for U^g and U^b that satisfy the two first-order conditions of that problem.
 - (b) Given U^g and U^b and $\{e, \theta\}$, which are enumerated, back out U^e from (25), and then back out c using U^g , U^b , e , and U^e from (26).
 - (c) Compute the planner's value that is associated with each effort level.³⁶

³³Since effort is bounded from below by zero, in general the first-order condition in equation (24) should not necessarily hold with equality. However, the first-order condition that I use over the finite grid of effort is only used for positive values of effort on the grid. Whenever the planner considers assigning a zero effort to the agent, the problem is simplified to be of full insurance (for which I use a closed-form exact value), implying that in this case the first-order condition is not used. Notice that in my quantitative analysis the optimal effort is always strictly positive (Figure 7). While visually the effort (top panel) is very low in this figure, it is positive as can be inferred from the positive job-finding probability (bottom figure), which can only happen with positive effort; see equation (6).

³⁴The following expression is also a function of θ and e but recall that we enumerate on those two decision variables.

³⁵The algorithm for updating W is similar to the one of updating V described below, but simpler as W includes less decision variables than V .

³⁶Notice that for this stage I need to approximate the value function as the continuation values are, in general, not located on the grid of promised utility. For this approximation, I use linear interpolations (and extrapolations when appropriate) over the continuation values.

(iv) Choose $\forall U$ the $\{e, \theta\}$ that maximize V and replace the previous guess for V with that value.

(v) Repeat the stages above till V converges.

This algorithm is implemented using a uniform grid over 200 points over promised utility levels, $U \in [-625, -275]$. The corresponding monthly consumption levels for those values—from about \$3800 to \$8000—are illustrated in the horizontal axis of Figure 3.

APPENDIX C: FAKING SEARCH BEHAVIOR

This Appendix provides additional details on the extension of the model where the worker can fake her search behavior. It includes two parts: Section C.1 provides the details of the calibration. Section C.2 shows that Proposition 1 holds for preferences that are very close to the knife-edge case of coefficient of relative risk aversion of 0.5.

C.1 Calibration of the one-period model

The following parameters are common to all the preferences used in the analysis. The job-finding probability conditional on exerting job-search effort is 0.3. For promised utility, I choose a grid that ranges from $u(\underline{c})$ to $u(\bar{c})$, where $u(\underline{c}) = 2$ and $u(\bar{c}) = 10$.³⁷ Notice that for different preferences the utility grid is different. For comparison across different preferences, I use the grid over consumption rather than for utility.

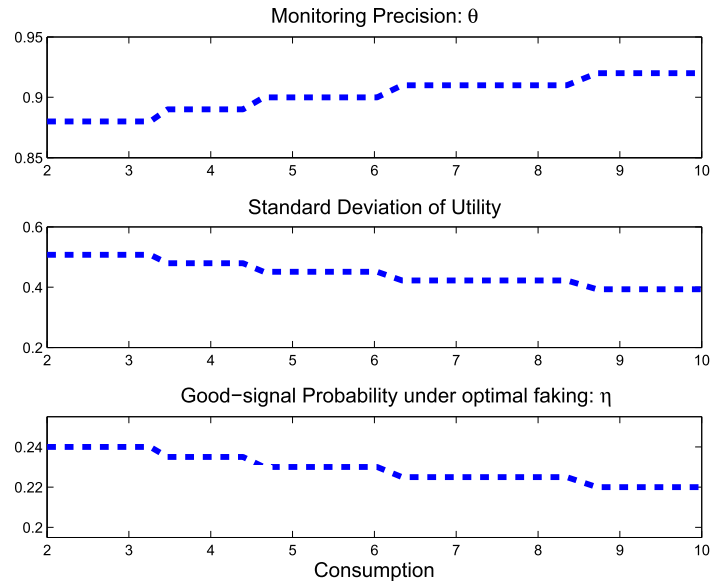


FIGURE 15. Optimal choices with and without faking search effort for $\sigma = 0.55$.

³⁷I choose $u(\underline{c}) > 0$ to avoid corner solutions.

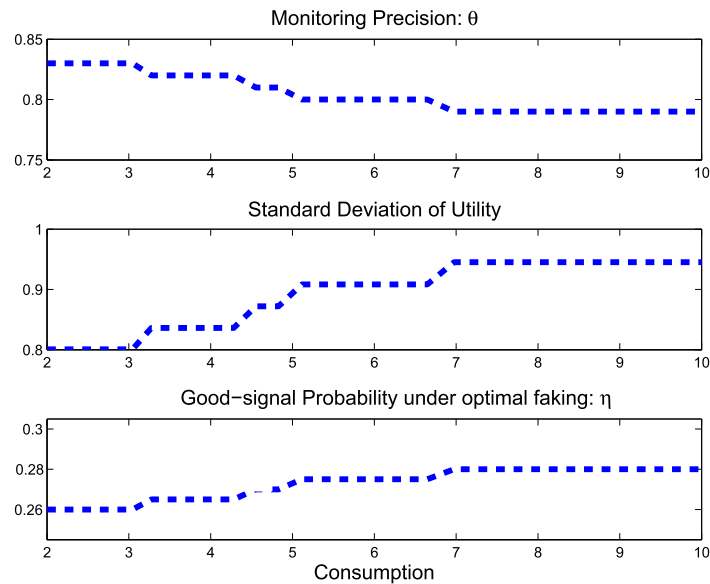


FIGURE 16. Optimal choices with and without faking search effort for $\sigma = 0.45$.

The remaining parameters are preferences specific. Disutility from effort, e , is given in utility terms and, therefore, is adjusted by preferences as the parameter that makes the worker indifferent between being employed and receiving a wage w and being unemployed and actively searching for a job while receiving a fraction b of her wage. The equation that determines this condition is: $u(w) = u(bw) - e$. As b is interpreted as UI benefits its value is set to 0.5. This is evaluated at the middle of the consumption support described above, that is, at $w = 6$. The values of e for CRRA preferences with a coefficient of $\{0.25, 0.50, 0.75\}$, are $\{2.072, 1.434, 0.996\}$, respectively.

The cost of providing monitoring with precision θ is $k\theta^{1+\eta}$, where $\eta = 0.4$ and k is equal to $\{0.4, 1.2, 2.0\}$ for CRRA preferences with a coefficient of $\{0.25, 0.50, 0.75\}$, respectively. The cost for the worker of creating a good signal conditional on not exerting job-search effort is $\zeta\eta^\psi$ where ζ is equal to four times the disutility cost of the job-search effort (e), and $\psi = 2$.

The parameters for the additional figures shown in C.2 for CRRA with coefficients of 0.55 and 0.45 are the same as those for CRRA with 0.50 described above.

C.2 Additional figures

In Proposition 1, I showed that without the ability of the worker to fake her search behavior, the monitoring precision increases and the spread decreases with promised utility if and only if the derivative of the inverse of the utility of the worker is convex. In this Appendix, I show that this result holds quantitatively for the model where the worker can fake her search behavior even as getting very near to the knife-edge case of CRRA with a coefficient of 0.5. The following figures show the results for coefficients of 0.55

(derivative of the inverse is strictly convex) and 0.45 (derivative of the inverse is strictly concave).

REFERENCES

- Abbring, J. H., G. J. van den Berg, and J. C. van Ours (2005), “The effect of unemployment insurance sanctions on the transition rate from unemployment to employment.” *Economic Journal*, 115, 602–630. [696]
- Abdulkadiroglu, A., B. Kuruscu, and A. Şahin (2002), “Unemployment insurance and the role of self-insurance.” *Review of Economic Dynamics*, 5 (3), 681–703. [723]
- Aiyagari, R. S. and F. Alvarez (1995), “Efficient dynamic monitoring of unemployment insurance claims.” Working paper. [696]
- Ashenfelter, O., D. Ashmore, and O. Deschenes (2005), “Do unemployment insurance recipients actively seek work? Evidence from randomized trials in four U.S. states.” *Journal of Econometrics*, 125 (1–2), 53–75. [708]
- Atkeson, A. and R. E. J. Lucas (1995), “Efficiency and equality in a simple model of efficient unemployment insurance.” *Journal of Economic Theory*, 66 (1), 64–88. [696]
- Attanasio, O., H. Low, and V. Sanchez-Marcos (2008), “Explaining changes in female labor supply in a life-cycle model.” *American Economic Review*, 98 (4), 1517–1552. [704]
- Becker, G. S. (1968), “Crime and punishment: An economic approach.” *Journal of Political Economy*, 76, 169. [696]
- Boone, J., P. Fredriksson, B. Holmlund, and J. C. van Ours (2007), “Optimal unemployment insurance with monitoring and sanctions.” *Economic Journal*, 117 (518), 399–421. [697, 718]
- Cogan, J. F. (1981), “Fixed costs and labor supply.” *Econometrica*, 49 (4), 945–963. [704]
- Corson, W. and W. Nicholson (1985), “Evaluation of the charlseton ClaimantPlacement and work test demonstration.” Unemployment Insurance Occasional Papers 85-2. US Department of Labor. [708]
- DOL (2006), “UI reforms CY 2004 annual report.” Department of Labor, U.S. Bureau of Labor Statistics. [708]
- DOL (2015), “National compensation survey: Occupational wages in the United States.” Department of Labor, U.S. Bureau of Labor Statistics. [704]
- Eckstein, Z. and K. I. Wolpin (1989), “Dynamic labour force participation of married women and endogenous work experience.” *Review of Economic Studies*, 56 (3), 375–390. [704]
- Elsby, M. W. L., B. Hobijn, and A. Şahin (2010), “The labor market in the great recession.” *Brookings Papers on Economic Activity*, 41 (1), 1–69. [706]
- Fougere, D., J. Pradel, and M. Roger (2009), “Does the public employment service affect search effort and outcomes?” *European Economic Review*, 53 (7), 846–869. [695]

- Fredriksson, P. and B. Holmlund (2006a), "Improving incentives in unemployment insurance: A review of recent research." *Journal of Economic Surveys*, 20 (3), 357–386. [697]
- Fredriksson, P. and B. Holmlund (2006b), "Optimal unemployment insurance design: Time limits, monitoring, or workfare?" *International Tax and Public Finance*, 13 (5), 565–585. [697]
- Fuller, D. L., B. Ravikumar, and Y. Zhang (2015), "Unemployment insurance fraud and optimal monitoring." *American Economic Journal: Macroeconomics, American Economic Association*, 7 (2), 249–290. [697]
- Grubb, D. (2000), "Eligibility criteria for unemployment benefits." OECD Economic Studies, 31. [693, 696, 707, 724]
- Hausman, J. A. (1980), "The effect of wages, taxes, and fixed costs on women's labor force participation." *Journal of Public Economics*, 14 (2), 161–194. [704]
- Hopenhayn, H. A. and J. P. Nicolini (1997), "Optimal unemployment insurance." *Journal of Political Economy*, 105 (2), 412–438. [694, 697, 715, 717]
- Hopenhayn, H. A. and J. P. Nicolini (2009), "Optimal unemployment insurance and employment history." *Review of Economic Studies*, 76 (3), 1049–1070. [727]
- Johnson, T. R. and D. H. Klepinger (1994), "Experimental evidence of unemployment work-search policies." *Journal of Human Resources*, 29 (3), 695–717. [695]
- Klepinger, D. H., T. R. Johnson, J. M. Joesch, and J. M. Benus (1997), "Evaluation of the Maryland unemployment insurance work search demonstration." Maryland Department of Labor, Licensing and Regulation. [695]
- Krueger, A. B. and A. Mueller (2010), "Job search and unemployment insurance: New evidence from time use data." *Journal of Public Economics*, 94 (3–4), 298–307. [706, 707]
- Krueger, A. B. and A. I. Mueller (2012), "The lot of the unemployed: A time use perspective." *Journal of the European Economic Association*, 10 (4), 765–794. [707]
- Lalive, R., J. C. van Ours, and J. Zweimüller (2005), "The effect of benefit sanctions on the duration of unemployment." *Journal of the European Economic Association*, 3 (6), 1386–1417. [696]
- Meyer, B. D. (1995), "Lessons from the U.S. unemployment insurance experiments." *Journal of Economic Literature*, 33 (1), 91–131. [708]
- MFIP (2000), "Minnesota family investment program. Program evaluation report: Welfare reform." REPORT 00-03. Chapter 4: Employment services for MFIP participants. Office of the Legislative Auditor. State of Minnesota., 102. [709]
- Paserman, M. D. (2008), "Job search and hyperbolic discounting: Structural estimation and policy evaluation." *Economic Journal*, 118 (531), 1418–1452. [705]
- Pavoni, N. (2009), "Optimal unemployment insurance, with human capital depreciation and duration dependence." *International Economic Review*, 50 (2), 323–362. [723]

Pavoni, N. and G. Violante (2007), “Optimal welfare-to-work programs.” *Review of Economic Studies*, 74 (1), 283–318. [697, 699]

Popov, L. (2009), “Stochastic costly state verification and dynamic contracts.” Report, University of Virginia. [696]

Ravikumar, B. and Y. Zhang (2012), “Optimal auditing and insurance in a dynamic model of tax compliance.” *Theoretical Economics*, 7 (2), 241–282. [696]

Shimer, R. (2012), “Reassessing the ins and outs of unemployment.” *Review of Economic Dynamics*, 15 (2), 127–148. [706]

Shimer, R. and I. Werning (2008), “Liquidity and insurance for the unemployed.” *American Economic Review*, 98 (5), 1922–1942. [723]

Spear, S. E. and S. Srivastava (1987), “On repeated moral hazard with discounting.” *Review of Economic Studies*, 54 (4), 599–617. [703]

van den Berg, G. J. (1994), “The effects of the job offer arrival rate on the duration of unemployment.” *Journal of Labor Economics*, 12 (3), 478–498. [695]

van den Berg, G. J., B. van der Klaauw, and J. van Ours (2004), “Punitive sanctions and the transition rate from welfare to work.” *Journal of Labor Economics*, 22 (1), 211–241. [695, 696]

van den Berg, G. J. and B. van der Klaauw (2006), “Counseling and monitoring of unemployed workers: Theory and evidence from a controlled social experiment.” *International Economic Review*, 47 (3), 895–936. [695]

Vroman, W. and S. A. Woodbury (2001), *The Incidence and Cost of Wrongfully Denied Unemployment Benefits*. Technical report. Available at <http://wdr.doleta.gov/conference/pdf/vroman.pdf>. [708]

Wang, C. and S. D. Williamson (2002), “Moral hazard, optimal unemployment insurance, and experience rating.” *Journal of Monetary Economics*, 49 (7), 1337–1371. [698]

Woodbury, S. A. (2002), “Unemployment insurance overpayments and underpayments.” Unemployment Fraud and Abuse, Hearing before the Subcommittee on Human Resources of the Committee on Ways and Means, House of Representatives. Available at <http://research.upjohn.org/testimonies/10>. [708]

Co-editor Karl Schmedders handled this manuscript.

Manuscript received 13 April, 2015; final version accepted 31 January, 2018; available online 21 May, 2018.