Inconsistent Time Preferences and On-the-job Search – When it Pays to be Naive *

Matthias Fahn[†] & Regina Seibel[‡]

May 22, 2020

Abstract

We study optimal employment contracts for present-biased employees who can conduct on-the-job search. Presuming that firms cannot offer long-term contracts, we find that individuals who are naive about their present bias will actually be better off than sophisticated or time-consistent individuals. They also search more, which partially counteracts the inefficiencies caused by their present bias. Moreover, firms might benefit from being ignorant of the extent of an employee's naiveté. Our results also indicate that naive employees might be harmed by policies such as a minimum wage, whereas sophisticated employees are (weakly) better off.

JEL Codes: D21, D83, D90, J31, J32
Keywords: Present bias, on-the-job search

---

*We are grateful to Florian Englmaier, Paul Heidhues, Takeshi Murooka, Nick Netzer, Armin Schmutzler, Heiner Schumacher, Peter Schwardmann, Ferdinand von Siemens, as well as seminar participants at the COPE 2018 (Munich) and the annual meeting of the Verein für Socialpolitik 2018 (Freiburg) for valuable comments and suggestions. Financial support by Deutsche Forschungsgemeinschaft through CRC TRR 190 is gratefully acknowledged.
[†]JKU Linz & CESifo; matthias.fahn@jku.at
[‡]University of Zurich; regina.seibel@econ.uzh.ch

1 Introduction

People suffer from self-control problems which are often caused by inconsistent time preferences. A huge literature has explored how firms lure consumers into inefficient "exploitative contracts" and thereby extract substantial rents from those who are naive about their present bias.[1] However, although recent observations suggest that time-inconsistent preferences matter in the workplace as well (Kaur et al., 2010, 2015), evidence that firms also try to attract and exploit naive *employees* remains scarce.[2] Naturally, a (so far) limited amount of evidence does not implicate the absence of exploitative contracts in the workplace. Nonetheless, wrong expectations concerning one's future behavior might have different consequences for employees than for consumers because employment relationship are inherently incomplete and include dimensions beyond a mere exchange of services.

In this paper, we show that the misperception of one's future behavior does not necessarily harm employed invidividuals who conduct on-the-job search, namely if firms cannot commit to long-term contracts. Whereas employees with inconsistent time preferences generally search too little from the perspective of earlier periods, those who are naive about their present bias actually *search more* and are *better off* than sophisticated employees. As a consequence, firms' profits are lower when hiring individuals they know to be naive. Moreover, being ignorant about an employee's naiveté can increase a firm's profits. If it does not, firms might completely abstain from hiring naive employees.

Indeed, large numbers of job-to-job transitions indicate that on-the-job search is a significant force behind labor market dynamics.[3] At the same time, search activities on labor markets are mostly perceived to be caused by information frictions which prevent an immediate matching of workers with their optimal job types. There, heterogeneities of workers and jobs have gained considerable attention as main drivers of these frictions.[4] But less focus has been put on how the trade-off between costly search effort today and potential benefits later on determines the extent of the generated inefficiencies. Exceptions are DellaVigna and Paserman (2005) and Paserman (2008) who, among others, have recently incorporated "behavioral" assumptions into job search models. As a survey by Cooper and Kuhn (forthcoming) indicates, this approach can contribute to a better understanding of outcomes observed in labor markets.

[1]See DellaVigna (2009), Kőszegi (2014), or Heidhues and Kőszegi (2018) for overviews.

[2]Exceptions are Bubb and Warren (forthcoming) and Hoffman and Burks (forthcoming).

[3]For example, Bjelland et al. (2011) find that employer-to-employer flows accounted for around 4% of total employment in the US between 1991 and 2003; see Fallick and Fleischman (2001) or Nagypál (2008) for further evidence.

[4]See Pissarides (1994), Mortensen (2000), Moscarini (2005), or Gautier et al. (2010).

In this paper, we explore how inconsistent time preferences affect on-the-job search, i.e., search behavior of the employed. We develop a three-period model in which a principal hires an agent. The agent receives a wage and can conduct on-the-job search. Wages are determined in every period, with the principal making take-it-or-leave-it offers but being unable to commit to long-term contracts.[5] Now, a higher intensity of on-the-job search yields higher search costs for the agent, but also increases the likelihood of receiving an attractive job offer in the subsequent period. Following arguments developed by Pissarides (1992) or Nagypál (2005), and evidence presented by Biewen and Steffes (2010), Mueller (2010), or Cingano and Rosolia (2012), we assume that on-the-job search is more effective than search out of unemployment. Therefore, on-the-job search can also be regarded as a non-pecuniary benefit of employment, which allows for a negative wage premium that pushes the agent's compensation below his outside option.

The agent has a time-inconsistent taste for immediate gratification. Thus, the agent's search effort is too low from the perspective of earlier periods. This relates to results delivered by DellaVigna and Paserman (2005) and Paserman (2008), who show that search effort is significantly reduced if individuals are time-inconsistent. Whereas DellaVigna and Paserman (2005) and Paserman (2008) only consider job search out of unemployment, Cho and Lewis (2011) provide indicative evidence that a present bias also decreases on-the-job search. They observe a substantial gap between turnover intentions and turnover behavior among employees.

Moreover, the agent can either be sophisticated or naive (Laibson, 1997, O'Donoghue and Rabin, 1999b). We first assume that the principal can observe the extent of the agent's naiveté, for example because of personality tests conducted during the hiring process, and later consider the case of asymmetric information. Whereas the sophisticated agent perfectly anticipates his future present bias, the naive agent expects to be time-consistent later on. This affects the size of feasible wage cuts to exploit the benefits of on-the-job search. The period-2 wage completely extracts the agent's expected net benefits from period-2 search. But, due to the agent's time-inconsistency, discounting between periods 2 and 3 is stronger from the perspective of the second than from the perspective of the first period. From the perspective of period 1, the agent's utility from period-2 search thus is strictly positive, even taking into account the period-2 wage. But this is only fully recognized by a so-

[5]Note that this assumption differs from many contributions to the job-search literature, where long-term commitment by firms is mostly assumed. In a recent contribution, though, Board and Meyer-Ter-Vehn (2015) rule out such long-term commitment in a model of on-the-job search. They are able to generate a number of results that are consistent with empirical observations, for example that workers' retentation rates, motivation, and productivities are higher at high-wage firms. We further discuss this assumption in Section 6.1.

phisticated agent who is hence willing to accept an additional wage reduction in the first period. In contrast, a naive agent does not anticipate his future present bias and consequently does not perceive his period-2 utility to be positive. Therefore, he only accepts a wage reduction that reflects the benefits from period-1 search. All this implies that the *realized* utility of a naive agent at the beginning of period 1 exceeds the utility of a sophisticated agent. This also has implications for period-1 search effort. Since the period-1 wage is sunk when selecting search effort, a sophisticated agent's perceived utility from staying with the current employer is higher than a naive agent's. Therefore, the latter sees a higher net benefit from receiving an outside offer in period 2, which lets him search more extensively. If on-the-job search indeed reduces matching frictions (what the huge extent of job-to-job flows as described by Fallick and Fleischman, 2001, Nagypál, 2008, or Bjelland et al., 2011 indicates), a more pronounced naiveté of employees might then increase the efficiency of labor markets

Now, the naive agent's higher wage and search effort imply that the principal's profits with a sophisticated agent are larger. This stands in stark contrast to most of the literature on present-biased preferences, where naive individuals generally are worse off and generate less efficient outcomes than those who are sophisticated. Firms design "exploitative contracts" to actively attract consumers they expect to mispredict their own future use of a purchased product and charge high prices when agents change their plans.[6] Our paper indicates that these results cannot necessarily be extended to employment relationships which entail many trade-offs not present in customer-firm relationships.[7] Further support for this argument is provided in Section 4, where we consider asymmetric information in the sense that the principal is not able to observe whether the agent is naive or sophisticated (as Eliaz and Spiegler, 2006, we assume that the agent's present bias is common knowledge). Different from previous research,[8] we find that the principal's profits when being ignorant can actually be larger than if she can observe the agent's naiveté, namely if the likelihood of facing a naive type is small. Hence, personality tests in the hiring process which allow firms to develop a better idea about their employees' characteristics (and are becoming more and more precise due to the rapid technological progress) might also have negative side effects. This result is again driven by the naive agent's misperception of period-2 outcomes. First note that a separation of types with both accepting the period-1 wage is never optimal for the principal. Thus, if the naive agent expects both types to still be around at

[6]See Heidhues and Kőszegi, 2010, or Kőszegi, 2014 for a survey.

[7]A related argument is provided by Weinschenk (2020) who shows that naive agents can be better off in team settings.

[8]Such as Eliaz and Spiegler (2006); Heidhues and Kőszegi (2010); Heidhues and Kőszegi (2017)

the beginning of the second period, he perceives the principal to select one of the following two options. Either, she might offer a higher wage both types are going to accept, or she might offer a lower wage only the naive type (who believes to be time-consistent then, thus search more and accept a larger wage cut) expects to take. If the share of naive agents is low, the first option seems optimal, which in turn lets the naive agent wrongly anticipate a period-2 rent and thus accept a lower period-1 wage than with symmetric information. In this case, the principal's asymmetric-info profits would indeed be higher. If the share of naive agents is larger, the second option seems optimal, and the principal will generally be harmed by asymmetric information. Then, she might offer a higher period-1 wage, and sophisticated agents benefit from the existence of naive agents.[9] Alternatively, she might make an offer that the naive agent potentially rejects.

We have argued that firms' relationships with naive employees might contain trade-offs that have not been considered so far. Since labor markets often are heavily regulated, it is also important to explore the consequences of such regulation on these trade-offs. In Section 5, we show that a minimum wage can harm the naive agent, if the minimum wage is below the payments he ends up receiving but above the lower wage he expects to be paid in the future. Then, he wrongfully anticipates a higher wage tomorrow and thus accepts a higher wage reduction today. This result indicates that a non-binding minimum wage might *reduce* wages; indeed, evidence for such negative spillover effects has been presented by Neumark et al. (2004), Stewart (2012), or Hirsch et al. (2015). Furthermore, a minimum wage always makes the sophisticated (weakly) better off.

Finally, in Section 6 we discuss the robustness of our results once some of the underyling assumptions are relaxed.

2 Model Setup

Environment, Technology & Contracts

There is one principal ("she") and one agent ("he") who are active in three periods, $t = 1, 2, 3$. At the beginning of every period, the principal can make a take-it-or-leave-it employment offer to the agent. This employment offer consists of a payment as well as the request to conduct a task that is valuable to the principal. There, we abstract from incentive problems and assume that, upon acceptance, the agent fulfills the task (for example because effort is verifiable), and define $w_t \in \mathbb{R}$ as the agent's period-t net utility from employment. In the following, we use the term *wage*

[9]Such as in, for example, Ispano and Schwardmann (2017).

when referring to w_t, however bear in mind that w_t not only contains the agent's compensation, but also potential costs of work effort (in contrast to search effort, as defined below).

If the agent rejects the offer, he consumes his outside option utility which is normalized to zero. Upon acceptance, he receives w_t but is also able to conduct on-the-job search. More precisely, the agent chooses his level of search effort, $s_t \in [0, 1]$ which is associated with search costs $s^2/2$. Moreover, s_t equals the probability with which the agent receives an outside job offer in the subsequent period $t + 1$. An outside job offer involves a net benefit of $B > 0$ for the agent, and the game ends after such an offer has been accepted. We assume $B < 1$ to make sure that search effort always is below 1. For simplicity, we also assume that B is independent of time, hence the counteroffer – if received – is equally attractive at the beginning of periods 2 and 3. This setup is consistent with a counteroffer involving a new short-term contract. Hence, job search improves the position of the agent and increases his outside option, but does not grant him a fixed wage.[10]

Following Board and Meyer-Ter-Vehn (2015), we assume that the principal has no commitment to offer long-term contracts (and discuss this assumption in Section 6.1). Furthermore, the agent is only able to conduct on-the-job search, but cannot search after rejecting the principal's offer and being unemployed. Allowing the agent to search while being unemployed would not affect our results, as long as the associated (marginal) search benefits would be smaller. Indeed, a number of reasons have been identified for why search by the unemployed might be less effective than on-the-job search: A social stigma effect (see Biewen and Steffes, 2010, for evidence), a missing network (see Cingano and Rosolia, 2012, for evidence), the decay of human capital (Pissarides, 1992), or a higher likelihood of job termination by unemployed individuals (Nagypál, 2005) may reduce their chances of receiving a job offer. Generally, Mueller (2010) provides evidence that job search is more effective when being employed.

Finally, we assume that the level of the agent's search effort is not verifiable, hence no contract can be based on s_t.

Preferences

The agent is risk neutral and discounts future costs and future utilities in a quasi-hyperbolic way according to Laibson (1997) and O'Donoghue and Rabin (1999a). Immediate utilities are not discounted. Utilities after t periods are discounted with a factor $\beta\,\delta^{\mathrm{t}}$, with $\beta \in (0, 1]$ and $\delta \in (0, 1)$. Hence, an agent's preferences are

[10] In Section 6.2, we relax a number of assumptions regarding search benefits.

dynamically inconsistent. This implies that, conditional on accepting the principal's offers, the agent's utility at the beginning of period $t = 1$ equals

$$U_1 = w_1 - \frac{1}{2}s_1^2 + \beta\delta \left\{ s_1 B + (1 - s_1) \left[w_2 - \frac{1}{2}s_2^2 + \delta \left(s_2 B + (1 - s_2) w_3 \right) \right] \right\}.$$

There, note that the agent will not engage in on-the-job search in period 3 since the game ends afterwards.

In case he has not received an outside job offer, the agent's utility at the beginning of period $t = 2$ equals

$$U_2 = w_2 - \frac{1}{2}s_2^2 + \beta\delta \left(s_2 B + (1 - s_2) w_3 \right).$$

A comparison between U_1 and U_2 reveals the agent's time inconsistency. Whereas discounting between periods 2 and 3 equals δ from the perspective of period 1, the effective discount factor falls to $\beta\delta$ if evaluated from the perspective of period 2. In Section 6.3 we discuss the implications of the present bias referring to *all* subsequent actions or outcomes. Then, upon receiving the period-t wage, the agent would already discount this period's search effort with β.

Finally, the agent's utility at the beginning of period 3 – conditional on not having received an outside job offer and accepting the principal's offer – equals

$$U_3 = w_3.$$

The principal is not present biased and discounts future payoffs with δ. If the agent accepts her offer and conducts the task, the principal enjoys a benefit $\pi > 0$. The principal's outside utility (which she consumes if the agent rejects her employment offer or receives and potentially accepts an outside offer) equals $\underline{\pi}$, with $\underline{\pi} < \pi$. $\underline{\pi}$ might include the possibility of finding a new agent, but also potential replacement costs. In the following, we assume $B > (1 + \delta)(\pi - \underline{\pi})$, hence making an eventually successful counteroffer is not optimal for the principal. Although successful counteroffers are observed in reality, the mere amount of observed turnover levels in labor markets (as described in the Introduction) indicates that many outside offers are indeed accepted.

Perceptions

We assume that the agent might be sophisticated or (fully) naive concerning his future present bias.[11] A naive agent expects his present bias to disappear and to discount the future exponentially from the next period on. In contrast, a sophisticated agent perfectly anticipates his future present bias and thus also correctly predicts his future behavior.

Concerning inter-player perceptions, we assume common knowledge about the principal's time preferences. Moreover, the principal is aware of the agent's present bias as well as whether he is naive or sophisticated (in Section 4, we assume that the principal cannot observe the extent of the agent's naiveté). However, whereas the principal anticipates potential contradictions between planned and realized actions, the agent thinks that the principal shares his own perception regarding his future preferences.

Equilibrium

Following O'Donoghue and Rabin (1999a) and Englmaier et al. (2016), our equilibrium concept is perception-perfect equilibrium. There, a player's strategy maximizes expected payoffs in all subgames, given one's present preferences, and given one's perceptions of one's own future behavior as well as of the others'. This equilibrium concept enables us to support strategies that are built on a naive agent's inconsistent beliefs.

3 Results

In the following, we solve for a perception-perfect equilibrium that maximizes the principal's profits. Since the principal cannot commit to long-term contracts, her profits are maximized at the beginning of every period, and we have to apply backwards induction to solve for equilibrium outcomes. Furthermore, in a profit-maximizing equilibrium wage payments are minimized in every period. We will start with the time-consistent agent as a benchmark case for optimal search effort and wage setting. We then characterize equilibria for sophisticated and fully naive agents separately (starting with a time-consistent agent as a benchmark) and subsequently compare the two outcomes.

[11] In Appendix B, we show that our results are robust to allowing for partial naiveté.

Benchmark: Time-Consistent Agent

As a benchmark, we will first derive outcomes for a time-consistent agent (which is equivalent to setting $\beta = 1$). We start analyzing the third and last period, conditional on the agent not having received an outside job offer before. In $t = 3$, the agent will not search, as there is no period thereafter in which he could collect potential search benefits. Furthermore, the principal will offer the lowest wage such that the agent just accepts an employment offer. Therefore, the agent receives (and accepts) a wage offer $w_3^{TC} = 0$ in period 3, i.e., his net utility of being employed just equals his outside utility of zero.

In the period 2, conditional on not having received an outside job offer before, and conditional on having accepted the principal's employment offer, the agent chooses search effort to maximize $-s_2^2/2 + \delta s_2 B$, which yields a search level

$$s_2^{TC} = \delta B.$$

Since the agent's benefits from search, $-(s_2^{TC})^2/2 + \delta s_2^{TC} B = (\delta B)^2/2$, are strictly positive, and since the agent can only search if he is employed by the principal, the period-2 wage – the lowest wage still accepted by the agent – equals

$$w_2^{TC} = \frac{1}{2}(s_2^{TC})^2 - \delta s_2^{TC} B = -\frac{1}{2}(\delta B)^2 < 0.$$

Thus, on-the-job search can be regarded as a non-pecuniary benefit of being employed that allows the principal to reduce the period-2 wage below the agent's reservation utility (as previously derived by Board and Meyer-Ter-Vehn, 2015). At the beginning of period 2, taking into account w_2^{TC} and expected search benefits, the agent's utility equals his reservation utility of zero. Hence, the situation in the first period is equivalent to the second period, which implies that outcomes coincide as well. Lemma 1 collects the results for the benchmark case of a time-consistent agent.

Lemma 1 *A time consistent agent*

- *exerts the same search effort in periods 1 and 2, i.e.* $s_1^{TC} = s_2^{TC}$
- *receives the same wage in periods 1 and 2, i.e.* $w_1^{TC} = w_2^{TC} < 0$*. These wages are equal to the respective period's negative search benefit.*

The proof can be found in Appendix A.1.

3.1 Sophisticated Agent

Now, we analyze outcomes for a sophisticated agent and start with the third period, conditional on the agent not having received an outside job offer before. For the same reasons as for a time-consistent agent, $s_3^S = w_3^S = 0$.

In the second period, conditional on not having received an outside job offer before, the agent (having accepted the principal's employment offer) chooses search effort to maximize $-s_2^2/2 + \beta\delta s_2 B$, which yields a search level

$$s_2^S = \beta\delta B.$$

The period-2 wage w_2^S takes into account that search is only possible for the agent if being employed, and is set to satisfy $U_2^S = w_2^S - (s_2^S)^2/2 + \beta\delta\left(s_2^S B + (1 - s_2^S) w_3^S\right) = 0$. Thus,

$$w_2^S = \frac{1}{2}(s_2^S)^2 - \beta\delta s_2^S B = -\frac{1}{2}(\beta\delta B)^2 < 0.$$

As with the time-consistent agent, on-the-job search can be regarded as a non-pecuniary benefit of being employed that allows the principal to reduce the period-2 wage below the agent's reservation utility. Moreover, the agent's time-inconsistency gives the principal additional, intertemporal, opportunities to reduce wages.

Lemma 2 *Assume the agent is sophisticated. Then,*

- *search in the first period is lower than in the second period, i.e.* $s_1^S < s_2^S$
- *the period-1 wage is* lower *than the period-1 negative search benefit, i.e.* $w_1^S < \frac{1}{2}(s_1^S)^2 - \beta\delta s_1^S B$.

The proof can be found in Appendix A.2.

From the perspective of period 1, discounting between periods 2 and 3 amounts to δ, whereas the discount factor from the perspective of period 2 equals $\beta\delta$. This changes the relative assessment of costs and benefits of period-2 search. Thus, although w_2^S fully extracts the agent's net utility from search in period 2, it does so only *from the perspective of period 2*. But due to his present bias, the agent's period-2 search benefit in relation to his search costs is higher from the perspective of period 1. Plugging $w_3^S = 0$ and $w_2^S = \frac{1}{2}(s_2^S)^2 - \beta\delta s_2^S B$ into the agent's period-1 utility yields

$$U_1^S = w_1 - \frac{1}{2}(s_1)^2 + \beta\delta\left[s_1 B + (1 - s_1)\delta s_2^S B\,(1 - \beta)\right].$$

There, the last term, $(1 - s_1)\delta s_2^S B\,(1 - \beta)$, captures the "extra" utility of period-2 search when assessed from the perspective of earlier periods.

This yields two implications for period-1 outcomes. First, w_1^S is not only reduced by period-1 search benefits, but also by the agent's "extra" period-2 search benefits. Second, because the agent only enjoys these future search benefits if he continues to stay employed by the principal, his incentives to conduct on-the-job search are reduced in comparison to period 2.

Finally, note that, from the perspective of period 1, the agent searches "too little" for his own taste in period 2 (δB versus $\beta\delta B$). This confirms that the results DellaVigna and Paserman (2005) and Paserman (2008) have derived for search out of unemployment also hold for on-the-job search.

3.2 Naive Agent

Now, we assume that the agent is naive about his present bias. As before, the period-3 wage of the naive agent equals $w_3^N = 0$. Furthermore, upon not having received an outside job offer and having accepted the principal's employment offer, the naive agent's effective search effort in period $t = 2$ also maximizes $-(s_2)^2/2 + \beta\delta s_2 B$, yielding a search level

$$s_2^N = \beta\delta B.$$

Furthermore,

$$w_2^N = \frac{1}{2}(s_2^N)^2 - \beta\delta s_2^N B = -\frac{1}{2}(\beta\delta B)^2 < 0.$$

Whereas $w_2^N = w_2^S$ and $s_2^N = s_2^S$, the naive agent does not anticipate these outcomes in period 1. There, he expects to be an exponential discounter from period 2 on and therefore to maximize $-\frac{1}{2}(s_2)^2 + \delta s_2 B$. This implies that, from the perspective of period 1, the agent perceives to choose a search level

$$\tilde{s}_2^N = \delta B.$$

Because $\tilde{s}_2^N > s_2^N$, the agent overestimates his future search effort. As a consequence, in period 1 the naive agent *underestimates* his period-2 wage. He expects to be offered a wage $\tilde{w}_2^N = (\tilde{s}^N)^2/2 - \delta\tilde{s}^N B$ which is smaller than the period-2 wage he is effectively willing to accept, w_2^N.

The naive agent's behavior in $t = 1$ is thus determined by his perceptions of future outcomes, not their true realizations:

Lemma 3 *Assume the agent is naive. Then,*

- *search efforts in the first and second period are equal, i.e.* $s_1^N = s_2^N$

- *the period-1 wage is equal to the period-1 negative search benefit, i.e.* $w_1^N = \frac{1}{2}(s_1^N)^2 - \beta\delta s_1^N B$.

The proof can be found in Appendix A.3.

From the perspective of period $t = 1$, the naive agent expects to have a period-2 net utility of zero. The principal thus is not able to collect the additional search benefits that stem from the agent's time inconsistency.

3.3 Comparison

Now, we compare outcomes of a naive and a sophisticated agent. First, recall that $s_2^S = s_2^N$ as well as $w_2^S = w_2^N$ and $w_3^S = w_3^N$. Therefore, realized outcomes in periods two and three are identical. However, $s_2^S < \tilde{s}_2^N$ and $w_2^S > \tilde{w}_2^N$. This difference in anticipated behavior lets period-1 search effort and wages of a naive and a sophisticated agent differ.

Proposition 1 $s_1^N > s_1^S$, *i.e. the period-1 search effort of a naive agent is higher than of a sophisticated agent.*

The proof can be found in Appendix A.4.

From the perspective of period 1, a sophisticated agent perceives his period-2 net utility from staying with the principal to be positive, whereas a naive agent (wrongly) perceives it to be zero. Thus, the relative marginal benefits of obtaining an outside job offer are higher for the latter, who consequently searches more.

Next, we show that the naive is better off than the sophisticated agent. Thereby, we compare *realized* and not perceived utility levels. First, note that realized utility levels at the beginning of period 2 are the same for both types of agents, namely $U_2^S = U_2^N = 0$. In the first period, the sophisticated agent also has realized utility $U_1^S = 0$, whereas the naive agent only *perceives* his utility level to be $\tilde{U}_1^N = 0$. His realized utility, however, is higher.

Proposition 2 *The naive agent has a strictly positive realized period-1 utility,* $U_1^N > 0$, *and is consequently better off than the sophisticated agent.*

The proof can be found in Appendix A.5.

In periods 2 and 3, both types of agents exert the same search effort and end up getting the same wages. Only the naive agent's period-1 wage is larger (in relation to search benefits) than the wage of the sophisticated agent. The naive agent thus underestimates his total utility and will only later on recognize this unexpected rent.

Our results differ from much of the literature on inconsistent time preferences. There, naive consumers and/or employees are generally worse off than sophisticated ones. The reason is that firms design exploitative contracts where individuals pay high prices when changing their plans. Naive consumers' wrong perceptions of their future actions let these exploitative contracts seem attractive.[12] We argue that this picture might not be complete if firms are unable to commit to future contracts.

3.4 Principal

Now, we compare the principal's payoffs when employing a sophisticated agent to the case of employing a naive agent. Recall that $B > (1-\delta)(\pi - \underline{\pi})$, where π is the principal's per-period payoff from keeping the agent, and $\underline{\pi}$ her per-period payoff after losing the agent. Hence making a successful counteroffer would not be profitable.

From the second period onwards, naive and sophisticated agent are identical in terms of search effort and wages, thus the relative benefits to the principal are solely determined by period-1 outcomes.

Proposition 3 *The principal's profits with a sophisticated agent are higher than with a naive agent.*

The proof can be found in Appendix A.6.

The principal prefers to employ a sophisticated agent who receives a lower wage and conducts less search in the first period. The latter increases profits because $\pi > \underline{\pi}$.

4 Asymmetric Information

We have shown that the naive agent benefits from being naive, and that the principal's profits with a sophisticated agent are larger than with a naive agent. However, these results rely on the principal knowing the agent's type (or, more precisely, they rely on the naive agent's belief that the principal knows his type), which we now

[12] See Kőszegi (2014) for a survey on exploitative contracts in an IO context. Eliaz and Spiegler (2006), Gilpatric (2008), or Englmaier et al. (2016) analyze settings more related to ours, where firms exploit employees' misperceptions regarding their future behavior.

refer to as symmetric information.[13] In this section, we consider the case of asymmetric information in the sense that the principal is not able to observe the agent's naiveté, neither at the time of contracting nor at any later point in time. As in Eliaz and Spiegler (2006), we consider the level of β as common knowledge and focus on uncertainty about the agent's extent of naiveté. We show that the principal can benefit from being ignorant about the agent's type, in particular if the share of naive types is small. Otherwise, the principal is likely to be worse off than with symmetric information, but might abstain from hiring the naive agent at all.

To derive these results, we assume that the principal is randomly matched with an agent before the employment relationship starts and that the agent is naive with probability α_1 and sophisticated with probability $1-\alpha_1$ (we use a subscript because period-2 shares might differ due to different search levels in the first period; see the proof to Proposition 4). α_1 is common knowledge, thus the naive agent perceives a share α_1 of agents to be exponential discounters from period 2 on (like himself) and a share $1-\alpha_1$ to be time-inconsistent and sophisticated. The sophisticated agent knows that all agents are time-inconsistent. Finally, we assume that if the principal abstains from making an offer to the agent, or if the agent does not accept her offer, she cannot employ him in later periods.

Now, the symmetric-info result that realized outcomes in periods 2 and 3 solely rely on the agent's level of present bias β also extends to the present setting. Only the naive agent's period-1 perception regarding those outcomes might be different, which will influence his reservation wage and search in period 1.

In the following, we characterize potential outcomes, how they rely on the share of naive agents, α_1, and on players' beliefs. There, we will apply Bayes' rule whenever possible as in a Perfect Bayesian Equilibrium. In the proof to Proposition 4, we also make precise which off-path beliefs support which outcomes.

Proposition 4 *If α_1 is sufficiently small, the principal's profits are larger with asymmetric than with symmetric information. Otherwise, her profits can be strictly smaller. Moreover, there are levels of α_1 such that the principal might not employ the naive type.*

The proof can be found in Appendix A.7.

First, note that a separating contract in which agents accept different wages in the first period generally is not optimal. In the proof to Proposition 4, we show that

[13] Note that our previous analysis does not assume symmetric information in the strict sense, since the naive agent does not share the principal's belief about his own future preferences. This is a form of "non-common priors", as stated by Eliaz and Spiegler (2006).

if a separating equilibrium exists, a pooling equilibrium in which both types receive and accept the same period-1 wage always yields higher profits.

Second, the results rely on the naive agent's belief about the principal's period-2 offer, but also on the principal's belief about the naive's period-1 choice, as well as the naive's belief about the principal's belief regarding each type's acceptance decision: If both types are still around in the second period, the principal has two options (from the perspective of the naive agent's period-1 self). Either, she offers a high wage both types are going to accept. Or, she offers a low wage only the naive type expects to accept due to his perceived higher period-2 search effort. If α_1 is sufficiently small, the former seems optimal. Then, the naive type expects to be paid a higher period-2 wage than with symmetric information, and in turn is willing to also accept a lower wage in the first period. In this case, the principal benefits from asymmetric information because both, naive and sophisticated agent, are accepting the wage the sophisticated type would be paid with symmetric information. Moreover, the naive type searches less in the first period because his perceived period-2 utility of staying with the principal is higher.

If α_1 is higher, the naive agent expects the principal to offer a low period-2 wage (in case both types have accepted the period-1 contract) and thus exclude the sophisticated type. Then, he is only willing to accept the period-1 contract in case he is paid the higher symmetric-information wage. But this wage has to be paid to the sophisticated type as well, hence can only be optimal if the share of sophisticated agents is sufficiently small. In this case, the sophisticated type benefits from the presence of naive agents. Moreover, the principal is worse off than with symmetric information because both types' search efforts are the same, but the sophisticated has to be paid a higher wage.

Finally, if the share of naive agents is too large for a high perceived period-2 wage but too low for a high period-1 wage, it is a dominant strategy for the principal to offer a low wage in the first period, even if the naive type decides to reject such a contract. In this case, naive agents might not be hired at all and the principal be worse off than with symmetric information: Whereas her profits when facing the sophisticated agent are the same, she does not make any profits with a naive type.

In order to make a more precise prediction concerning the naive agent's behavior for these levels of α_1, we would have to impose additional assumptions about his beliefs. The reason is that if the naive agent rejects a period-1 offer with probability 1 and perceives the principal to know about this, he expects the principal to offer w_2^S in the second period. But then it would be optimal for the naive agent to accept the period-1 offer, in which case it would be optimal for the principal (from the naive's perspective) to offer the lower w_2^N in the second period, and so on. It is beyond the scope of this paper to define a potential equilibrium for this case, however note that this perceived inconsistency does not affect the principal's behavior for these values of α_1. In the first period, she will offer w_1^S (independent of her expections about the naive agent's behavior), in the second she will offer w_2^S.[14]

Finally, we could incorporate recent evidence that many humans are naive about their own present bias but correctly predict other naive agents' time-inconsistency (Fedyk, 2018). This would make naive agents even more prone to believe that a high period-2 wage is optimal for the principal and accept a lower period-1 wage. Then, naive agents would not only be harmed by the principal's asymmetric information, but also by their own sophistication regarding others' present bias.

5 Effects of a minimum wage

The same mechanism that potentially lets the naive agent be harmed by asymmetric information – expecting a higher period-2 wage and thus accepting a lower period-1 wage – can also be applied to generate new insights on the effects of a minimum wage.[15] Whereas a sufficiently high minimum wage in our setup benefits all agents, an intermediate level can actually harm a naive agent: Assume there is a minimum wage that exceeds the period-2 wage a naive agent expects to be paid but is below the wages he is paid in the first and second period. Then, a naive agent wrongfully anticipates a rent in the second period and consequently accepts a higher wage reduction in the first. In addition, his period-1 search effort goes down. A sophisticated agent, on the other hand, would at all levels (weakly) benefit from a minimum wage, and also search less if the minimum wage is binding in period 2. These results are collected in the following proposition.

Proposition 5 *Assume there is a minimum wage $\overline{w} \geq -\frac{1}{2}(\delta B)^2$. Then, a higher minimum wage (weakly) reduces search effort of both types. Moreover, it has the following effects on wages and payoffs:*

- *For the naive agent, there exists a $\overline{w}^* \leq -\frac{1}{2}(\beta\delta B)^2$ such that $dw_1^N/d\overline{w} < 0$ and $dU_1^N/d\overline{w} < 0$ if $\overline{w} \in \left[\tilde{w}_2^N, \overline{w}^*\right)$. If $\overline{w} \geq \overline{w}^*$, $dU_1^N/d\overline{w} \geq 0$.*
- *For the sophisticated agent, $dU_1^S/d\overline{w} \geq 0$ and $dw_1^S/d\overline{w}$, $dw_2^S/d\overline{w} \geq 0$.*

[14] In the proof to Proposition 4, we also argue that equilibria in mixed strategies would not deliver clearer results.

[15] Recall that the w in our setting captures the agent's net benefits of being employed by the principal; as long as there is a monotonic relationship between wages and these net benefits, the following results hold.

The proof can be found in Appendix A.8.

If $\overline{w} \geq -\frac{1}{2}(\delta B)^2$, the minimum wage exceeds the naive agent's perceived period-2 wage (for $\overline{w} < -\frac{1}{2}(\delta B)^2$, a minimum wage is irrelevant in our setting).[16] Then, he reduces his search effort because he expects a higher utility conditional on not receiving an outside offer. Importantly, this holds even for a minimum wage that effectively never binds, i.e., for $\overline{w} < -\frac{1}{2}(\beta\delta B)^2$. In this case, he can be harmed by a higher minimum wage, because his perceived (but not real) period-2 utility goes up and he is thus willing to accept a lower wage in the first period. Note that these results rely on B being unaffected by a minimum wage (which might be the case if the outside offer is sufficiently large for a minimum wage to be irrelevant). But our results would be qualitatively robust even if B increased in $\overline{w}$, as long is this increase was not too large.

Concluding, our results indicate that a non-binding minimum wage might have *negative* spillover effects on higher wages and might reduce search effort and consequently *turnover*. Regarding the former, previous literature has mostly focused on explanations for observed positive spillover wages, which include forces outside our model (such as firms wanting to preserve their wage distribution). However, there is also evidence that spillover effects can indeed be negative. For example, Stewart (2012) examines the consequences of the British minimum wage. He finds that the growth of wages slightly above the minimum wage generally is *smaller* than what would have been expected without a minimum wage. Neumark et al. (2004) observe that, although immediate spillover effects are positive, lagged effects are strongly negative. Finally, Hirsch et al. (2015) provide indicative evidence that wages above the minimum wage increase less strongly than they would have without the minimum wage. Regarding turnover, there is abundant evidence on negative effects, however a number of theoretical explanations can already yield explanations. These explanations only hold for *binding* minimum wages, though, whereas we would also predict negative turnover effects of a *non-binding* minimum wage. To the best of our knowledge, this aspect has so far not been explored empirically (some indicative evidence is provided by Hirsch et al., 2015, who find that the negative effects of a minimum wage on turnover are not necessarily increasing in compliance costs, which are a measure of the extent to which minimum wages bind), so we hope our ideas might contribute to a better understanding of the consequences of a minimum wage.

[16]Of course, real-world minimum wages are not negative. However, note that our argument only relies on the relation of the minimum wage to the agent's reservation utility which we have normalized to zero.

6 Robustness

In the following, we discuss the robustness of our results once we relax some assumptions.

6.1 Long-term commitment

The assumption that the principal is not able to offer a long-term contract is crucial for our results. If the principal were able to commit to future wages, the sophisticated agent would not be affected, whereas the naive agent would be harmed. Then, the principal could offer w_2^S also to the naive agent and thus exploit his false expectations about future search effort while avoiding false expectations about future wages (in this case, the naive agent would wrongly perceive his period-1 utility to be positive). However, we would argue that forces outside our model might also create costs of commitment, and that the extent of long-term commitment is limited. First, uncertainty regarding future prospects might reduce the benefits of such commitment. Assume there is a chance that the principal's profits from employing the agent might drop in any period (for example due to demand fluctuations or productivity shocks), and the principal would rather terminate the relationship in that case. Then, the principal has to trade off the reduced flexilibity of commitment with the possibility to exploit the naive agent. Thus, we would predict naive agents to search more and earn higher wages particularly in industries or countries where commitment is not possible, or where it is not optimal because of uncertain future prospects.

Second, even if we take into account that commitment to future wages often is possible, such commitment generally does not hold for arbitrarily long time periods. At some point, the principal has the opportunity to end the contract (and then negotiate the terms of a continuation). Thus, if we extended the model to include more than three periods, our results would continue to hold even with commitment, as long as such commitment did not extend to the agent's entire career. Alternatively, we could interpret a period in our model to cover longer phases of an employee's career.

Finally, note that the discussion in this section also implies that more restrictive employment protection can harm naive present-biased employees because such laws allow for more commitment of employers.

6.2 Search Benefits

We have assumed that the net benefit of obtaining an outside offer is some fixed level B. We have done so for expositional simplicity, however this assumption is less restrictive than it may seem. In Appendix A.9, we show that our main results continue to hold if, first, we explicitly model an outside offer as a new short-term contract with the possibility to engage in further job search; second, offers are drawn from an exogenous distribution; and third, the game still ends in period 3 when accepting an offer, thus the benefits of search are not constant over time.[17]

Under the new assumptions, the analysis for the sophisticated agent is a little more involved than before, still our main results continue to hold. He continues to search in the second period even if period-1 search has been successful (whether with the principal or a new employer). Assessing period-2 search from the perspective of the first period, he therefore anticipates "extra" utility, not only from search when today's search is not successful, but also from continued search when today's search turns out to be successful. [18] This decreases the benefit of successful relative to unsuccessful search and thus search effort. Moreover, it allows the principal to further reduce the wage of the sophisticated agent in the first period, capturing both types of extra utility. As before, $U_1^S = 0$.

The analysis of the naive agent remains qualitatively unchanged: In period 1, he overestimates his period-2 search, irrespective of having received a new offer and the resulting wage cut, not anticipating his time-inconsistency (and therefore not considering any extra-utility of the above types). Consequently, he searches more than the sophisticated agent in the first period and only accepts a wage granting him a positive rent, $U_1^N > 0$.

6.3 Different timing of search cost assessment

When the agent assesses a period's wage and search effort at the beginning of a period, both relate to the "present". However, evidence indicates that a present bias is only about very recent events (O'Donoghue and Rabin, 2015). For example, Augenblick (2017) shows that the β discounts consumption already a few hours away and in particular that consumption more than a few days away is not included in the "present" an individual is biased towards. We now discuss the robustness of our results if we incorporate these aspects into our model. To do so, we might split each period into two steps and discount period-t search with β already *at the beginning* of period t. For example, at the beginning of period 2 an agent's utility would be

[17] In addition, we may allow for successful counter-offers by the principal, but this assumption is inconsequential for the agent's search behavior. Therefore, the results in this section are unaffected.

[18] Note that the latter "extra" utility is smaller than the former because, conditional on already having received a good offer, the benefit of continued search and consequently the principal's possibility to reduce the wage are smaller.

$$w_2 + \beta \left[-\frac{1}{2} s_2^2 + \delta \left(s_2 B + (1 - s_2) w_3 \right) \right],$$

whereas at the time when the search decision is made his utility would amount to

$$-\frac{1}{2} s_2^2 + \beta\delta \left(s_2 B + (1 - s_2) w_3 \right).$$

Then, the naive agent would also overestimate his search in $t = 2$ when assessing the period's wage offer. He would expect to search $\tilde{s}_2 = \delta B$ (compared to his realized search $s_2 = \beta\delta B$) and accept a wage satisfying $w_2 + \beta \left[-\tilde{s}_2^2/2 + \delta \tilde{s}_2 B \right] = 0$. Thus, the naive agent's *realized* period-2 utility would be negative also from the perspective of period 2 (whereas the sophisticated agent would still only accept a wage that yields a non-negative utility). However, this would not necessarily imply that the naive agent now always was worse off than the sophisticated agent. The reason is that his period-2 utility would continue to be positive from the perspective of earlier periods: Conditional on not receiving an outside offer, the naive agent's realized period-2 utility from the perspective of period 1 would amount to

$$\beta\delta \left[w_2 - s_2^2/2 + \delta s_2 B \right] = \frac{\delta}{2} (\beta\delta B)^2 (1 - \beta) > 0.$$

This is positive because the lower discounting between periods 2 and 3 from the perspective of the first (compared to the second) period more than makes up for the exploitative wage. His perceived period-2 utility from the perspective of period 1 would continue to be zero, though, thus the principal could not exploit this rent.

In this case, the naive agent's realized utility at the beginning of his career could still be positive, for example if we assumed that he did not immediately search for a better opportunity when starting a job but wait for at least one period (i.e., if search was not feasible or optimal in $t = 1$). Equivalently, we might assume that search was possible at the beginning of period 1, but the associated benefits smaller than in later period, for example because of human capital accumulation. Moreover, in a model with more than three periods, the positive benefits of future search would further accumulate and also increase his realized utility.

7 Conclusion

We have shown that a firm's relationships with its employees might entail substantially different trade-offs than with its consumers. As a consequence, present-biased agents can benefit from being naive – in a situation where they conduct on-the-job search and firms cannot commit to long-term contracts. Finally, we have demonstrated that the principal might be better off when being ignorant about the extent of the agent's naiveté, and also benefit from a moderate minimum wage.

References

Augenblick, N. (2017). Short-term time discounting of unpleasant tasks. 6.3

Biewen, M. and Steffes, S. (2010). Unemployment persistence: Is there evidence for stigma effects? *Economics Letters*, 106:188–190. 1, 2

Bjelland, M., Fallick, Bruce, H. J., and McEntarfer, E. (2011). Employer-to-employer flows in the united states: Estimates using linked employer-employee data. *Journal of Business & Economic Statistics*, 29(4):493–505. 3, 1

Board, S. and Meyer-Ter-Vehn, M. (2015). Relational contracts in competitive labour markets. *The Review of Economic Studies*, 82(2):490–534. 5, 2, 3

Bubb, R. and Warren, P. L. (forthcoming). An equilibrium theory of retirement plan design. *American Economic Journal: Economic Policy*. 2

Cho, Y. J. and Lewis, G. B. (2011). Turnover intention and turnover behavior. *Review of Public Personnel Administration*, 32(1):4–23. 1

Cingano, F. and Rosolia, A. (2012). People i know: Job search and social networks. *Journal of Labor Economics*, 30(2):291–332. 1, 2

Cooper, M. and Kuhn, P. Behavioral job search. Handbook of Labor, Human Resources and Population Economics. Springer Nature. 1

DellaVigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic literature*, 47(2):315–72. 1, A.9

DellaVigna, S. and Paserman, M. D. (2005). Job search and impatience. *Journal of Labor Economics*, 23(3):527–588. 1, 3.1

Eliaz, K. and Spiegler, R. (2006). Contracting with diversely naive agents. *The Review of Economic Studies*, 73(3):689. 1, 8, 12, 4, 13, B.1

Englmaier, F., Fahn, M., and Schwarz, M. A. (2016). Long-term employment relations when agents are present biased. 2, 12

Fallick, B. C. and Fleischman, C. A. (2001). The importance of employer-to-employer flows in the u.s. labor market. 3, 1

Fedyk, A. (2018). Asymmetric naivete: Beliefs about self-control. *Available at SSRN 2727499*. 4

Gautier, P. A., Teulings, C. N., and Van Vuuren, A. (2010). On-the-job search, mismatch and efficiency. *Review of Economic Studies*, 77(1):245–272. 4

Gilpatric, S. M. (2008). Present-biased preferences, self-awareness and shirking. *Journal of Economic Behavior & Organization*, 67(3):735–754. 12

Heidhues, P. and Kőszegi, B. (2010). Exploiting naivete about self-control in the credit market. *American Economic Review*, 100(5):2279–2303. 6, 8

Heidhues, P. and Kőszegi, B. (2017). Naïveté-based discrimination. *The Quarterly Journal of Economics*, 132(2):1019–1054. 8

Heidhues, P. and Kőszegi, B. (2018). Behavioral industrial organization. In Bernheim, B. D., DellaVigna, S., and Laibson, D., editors, *Handbook of Behavioral Economics - Foundations and Applications 1*, volume 1 of *Handbook of Behavioral Economics: Applications and Foundations 1*, pages 517 – 612. North-Holland. 1

Hirsch, B. T., Kaufman, B. E., and Zelenska, T. (2015). Minimum wage channels of adjustment. *Industrial Relations: A Journal of Economy and Society*, 54(2):199–239. 1, 5

Hoffman, M. and Burks, S. V. (forthcoming). Worker overconfidence: Field evidence and implications for employee turnover and firm profits. *Quantitative Economics*. 2

Ispano, A. and Schwardmann, P. (2017). Competition over cursed consumers. 9

Kaur, S., Kremer, M., and Mullainathan, S. (2010). Self-control and the development of work arrangements. *American Economic Review*, 100(2):624–28. 1

Kaur, S., Kremer, M., and Mullainathan, S. (2015). Self-control at work. *Journal of Political Economy*, 123(6):1227–1277. 1

Kőszegi, B. (2014). Behavioral contract theory. *Journal of Economic Literature*, 52(4):1075–1118. 1, 6, 12

Laibson, D. (1997). Golden eggs and hyperbolic discounting*. *The Quarterly Journal of Economics*, 112(2):443–478. 1, 2

Mortensen, D. T. (2000). Equilibrium unemployment with wage posting: Burdett-mortensen meets pissarides. 4

Moscarini, G. (2005). Job matching and the wage distribution. *Econometrica*, 73:481–516. 4

Mueller, A. (2010). On-the-job search and wage dispersion: New evidence from time use data. *Economics Letters*, 109(2):124–127. 1, 2

Nagypál, É. (2005). Amplification of labor market fluctuations: Why vacancies don't like to hire the unemployed? Technical report, Society for Economic Dynamics. 1, 2

Nagypál, É. (2008). Worker reallocation over the business cycle: The importance of employer-to-employer transitions. *Manuscript, Northwestern Univ.* 3, 1

Neumark, D., Schweitzer, M., and Wacher, W. (2004). Minimum Wage Effects throughout the Wage Distribution. *Journal of Human Resources*, 39(2):425–450. 1, 5

O'Donoghue, T. and Rabin, M. (1999a). Doing it now or later. *The American Economic Review*, 89(1):103–124. 2, 2

O'Donoghue, T. and Rabin, M. (1999b). Incentives for procrastinators. *The Quarterly Journal of Economics*, 114(3):769–816. 1

O'Donoghue, T. and Rabin, M. (2015). Present bias: Lessons learned and to be learned. *American Economic Review*, 105(5):273–79. 6.3

Paserman, M. D. (2008). Job search and hyperbolic discounting: Structural estimation and policy evaluation. *The Economic Journal*, 118(531):1418–1452. 1, 3.1

Pissarides, C. (1994). Search unemployment with on-the-job serach. *Review of Economic Studies*, 61:457–475. 4

Pissarides, C. A. (1992). Loss of skill during unemployment and the persistence of employment shocks. *Quarterly Journal of Economics*, 107:1371–1391. 1, 2

Stewart, M. B. (2012). Wage inequality, minimum wage effects, and spillovers. *Oxford Economic Papers*, 64(4):616–634. 1, 5

Weinschenk, P. (2020). On the benefits of time-inconsistent preferences. 7

A Appendix

A.1 Proof of Lemma 1.

Conditional on not having received an outside job offer before, and conditional on having accepted the principal's employment offer at the beginning of $t = 1$, the agent chooses search effort to maximize,

$$-\frac{1}{2}(s_1)^2 + \delta \left\{ s_1 B + (1 - s_1) \left[w_2^{TC} - \frac{1}{2}(s_2^{TC})^2 + \delta \left(s_2^{TC} B + (1 - s_2^{TC}) w_3^{TC} \right) \right] \right\},$$

however taking into account $w_3^{TC} = 0$ and $w_2^{TC} = \frac{1}{2}(s_2^{TC})^2 - \delta s_2^{TC} B$. Therefore, the problem boils down to maximizing $-\frac{1}{2}(s_1)^2 + \delta s_1 B$, yielding $s_1^{TC} = \delta B$, as well as $w_1^{TC} = \frac{1}{2}(s_1^{TC})^2 - \delta s_1^{TC} B < 0$, hence $s_1^{TC} = s_2^{TC}$ and $w_1^{TC} = w_2^{TC}$. ■

A.2 Proof of Lemma 2.

s_1^S maximizes

$$\begin{aligned} &-\frac{1}{2}s_1^2 + \beta\delta \left\{ s_1 B + (1 - s_1) \left[w_2 - \frac{1}{2}s_2^2 + \delta s_2 B \right] \right\} \\ &= -\frac{1}{2}s_1^2 + \beta\delta \left\{ s_1 B + (1 - s_1)(1 - \beta) \beta\delta^2 B^2 \right\}, \end{aligned}$$

hence

$$s_1^S = \beta\delta \left[B - \delta s_2^S B (1 - \beta) \right].$$

It follows that $s_1^S < s_2^S = \beta\delta B$.

w_1^S is set to satisfy $U_1^S = w_1^S - \frac{1}{2}(s_1^S)^2 + \beta\delta \left[s_1^S B + (1 - s_1^S)\delta s_2^S B (1 - \beta) \right] = 0$, hence

$w_1^S = \frac{1}{2}(s_1^S)^2 - \beta\delta \left[s_1^S B + (1 - s_1^S)\delta s_2^S B (1 - \beta) \right] < \frac{1}{2}(s_1^S)^2 - \beta\delta s_1^S B.$ ■

A.3 Proof of Lemma 3

A fully naive agent perceives his first-period utility to be

$$\tilde{U}_1^N = w_1 - \frac{1}{2}(s_1)^2 + \beta\delta \left\{ s_1 B + (1 - s_1) \left[\tilde{w}_2^N - \frac{1}{2}(\tilde{s}_2^N)^2 + \delta \left(\tilde{s}_2^N B + (1 - \tilde{s}_2^N) w_3^N \right) \right] \right\}.$$

Making use of $\tilde{w}_2^N = \frac{1}{2}(\tilde{s}_2^N)^2 - \delta \tilde{s}_2^N B$ and $w_3 = 0$, this becomes

$$\tilde{U}_1^N = w_1 - \frac{1}{2}(s_1)^2 + \beta\delta s_1 B.$$

The Lemma immediately follows. ■

A.4 Proof of Proposition 1.

As shown in Lemma 2, period-1 search effort of a sophisticated agent equals

$$s_1^S = \beta\delta B\left[1 - \delta s_2^S\,(1-\beta)\right] = 0.$$

As shown in Lemma 3, period-1 search effort of a fully naive agent equals

$$s_1^N = \beta\delta B.$$

Given $s_2^S > 0$, the Proposition immediately follows. ■

A.5 Proof of Proposition 2.

The naive agent's realized utility level in period $t = 1$ amounts to

$$\begin{aligned} U_1^N =& w_1^N - \frac{1}{2}(s_1^N)^2 \\ &+ \beta\delta\left\{s_1^N B + (1-s_1^N)\left[w_2^N - \frac{1}{2}(s_2^N)^2 + \delta\left(s_2^N B + (1-s_2^N)w_3^N\right)\right]\right\}. \end{aligned}$$

.

Taking into account $w_1^N = \frac{1}{2}(s_1^N)^2 - \beta\delta s_1^N B$, $w_2^N = \frac{1}{2}(s_2^N)^2 - \beta\delta s_2^N B$ and $w_3^N = 0$,

$$U_1^N = \beta\,(1-\beta)\,\delta^2(1-s_1^N)s_2^N B > 0.$$

■

A.6 Proof of Proposition 3.

Recall that first-period wages are

- $w_1^N = (s_1^N)^2 - \beta\delta s_1^N B$
- $w_1^S = (s_1^S)^2 - \beta\delta\left[s_1^S B + (1-s_1^S)\delta s_2^S B\,(1-\beta)\right]$.

Taking into account that $s_1^N = \beta\delta B$ and $s_1^S = \beta\delta B\,[1 - \delta^2\beta(1-\beta)B]$, those amount to

- $w_1^N = -\frac{1}{2}(\beta\delta B)^2$
- $w_1^S = -\beta^2\delta^2 B^2 \left[\frac{1}{2}\left(1-\beta\delta^2 B(1-\beta)\right)^2 + \delta(1-\beta)\right]$

Hence

$$
\begin{aligned}
w_1^N - w_1^S &= -\frac{1}{2}(\beta\delta B)^2 + \beta^2\delta^2 B^2 \left[\frac{1}{2}\left(1-\beta\delta^2 B(1-\beta)\right)^2 + \delta(1-\beta)\right] \\
&= \beta^2\delta^3 B^2(1-\beta)\left[1-\beta\delta B + \frac{1}{2}\beta^2\delta^3 B^2(1-\beta)\right] > 0 \text{ if } \beta\delta B < 1.
\end{aligned}
$$

There, note that $\beta\delta B < 1$ holds in order to always guarantee.

Furthermore, $s_1^S < s_1^N$ (see Proposition 1), hence a naive agent is more likely to receive a new job offer – which reduces the principal's profits because of $\underline{\pi} < \pi$.

■

A.7 Proof of Proposition 4.

Generally, outcomes depend on players' beliefs. The naive agent observes the principal's first-period offer and forms beliefs regarding second-period offers, but also regarding the sophisticated agent's actions as well as the principal's beliefs on both types' first period acceptance choices. Also the principal's beliefs on the naive agent's first-period choices matter. The sophisticated agent, on the other hand, anticipates that outcomes from the second-period onwards will be independent of first-period choices (the principal then offers the same wage as with symmetric information). Thus, his first-period choice only depends on the first-period offer. In the following, we characterize potential outcomes and apply Bayes' rule whenever possible. For off-path events, we discuss which beliefs support a certain outcome without making a claim on their validity.

In the following, we first characterize potential pooling equilibria in which both types are offered the same wage in period 1. At the end of this proof, we also derive a potential separating equilibrium with different first-period offers, which however will not be optimal for the principal.

Pooling Equilibrium As in the game with symmetric information, we use backward induction to analyze the game. In the third period, agents do not search anymore, hence receive a wage of zero. This is correctly anticipated by both types, hence we move on to the second period.

If both types accepted the first-period contract, the likelihood of facing the naive type at the beginning of the second period might differ from the corresponding likelihood in the first period, α_1. The reason is that both types might engage in different levels of search, hence face different probabilities of receiving an outside offer. Let us denote the second period probability of facing a naive type as α_2, which due to Bayes' rule equals

$$\alpha_2 = \frac{\alpha_1(1-s_1^N)}{\alpha_1(1-s_1^N)+(1-\alpha_1)(1-s_1^S)} \in [0,1].$$

There, s_1^S and s_1^N correspond to the respective type's first-period equilibrium search effort. Also note that $\alpha_2 = 0$ ($\alpha_2 = 1$) in an equilibrium in which only the sophisticated (naive) type accepts the first period offer.

However, because the naive agent might have wrong expections on the principal's second-period offer, he might also misperceive the sophisticated agent's first-period search effort. Thus, the naive agent's perception of the principal's second-period likelihood of facing a naive agent, which we denote by $\tilde{\alpha}_2$, might differ from the true probability α_2. Now, we explore the principal's optimal second-period decisions (from the perspective of the naive agent's first-period self). First, if the first-period wage is *not* accepted by one of the types, the naive agent expects the principal to subsequently act as if she had full information. Denoting the principal's second period profits in case the naive agent's expects only the sophisticated type to have remained by $\tilde{\Pi}_2^S$, and by $\tilde{\Pi}_2^N$ the respective profits if the naive agent thinks that only he will then have remained,

$$\begin{aligned}\tilde{\Pi}_2^S &= \pi - w_2^S + \delta\left(s_2^S\underline{\pi} + (1-s_2^S)\pi\right)\\ \tilde{\Pi}_2^N &= \pi - \tilde{w}_2^N + \delta\left(\tilde{s}_2^N\underline{\pi} + (1-\tilde{s}_2^N)\pi\right).\end{aligned}$$

There, wage and search effort correspond to their full-information values.

Second, if the naive agent thinks that both types have accepted the first period offer, he perceives the principal to face the following trade-off. Either, the principal might set a rather low wage $\tilde{w}_2^N$ only the naive agent himself accepts; or, she might set a higher wage w_2^S that is satisfactory for both types (again, $\tilde{w}_2^N$ and w_2^S are equivalent to their full-information values). In the latter case, the naive agent perceives the principal's second-period profits to be

$$\tilde{\Pi}_2(w_2^S) = \tilde{\alpha}_2\tilde{\Pi}_2^N + (1-\tilde{\alpha}_2)\tilde{\Pi}_2^S - \tilde{\alpha}_2(w_2^S - \tilde{w}_2^N)$$

In the first case, those perceived profits amount to

$$\tilde{\Pi}_2(\tilde{w}_2^N) = \tilde{\alpha}_2\tilde{\Pi}_2^N + (1-\tilde{\alpha}_2)\underline{\pi}(1+\delta).$$

Hence, it would be optimal for the principal (from the perspective of the naive agent's first-period self) to employ both types of agents in the second period if $\tilde{\Pi}_2(\tilde{w}_2^N) \leq \tilde{\Pi}_2(w_2^S)$, i.e., if

$$\tilde{\alpha}_2 \leq \frac{\tilde{\Pi}_2^S - \underline{\pi}(1+\delta)}{\tilde{\Pi}_2^S - \underline{\pi}(1+\delta) + (w_2^S - \tilde{w}_2^N)} \equiv \bar{\alpha}_2.$$

Having computed the principal's profits and perceived optimal decisions for all possible cases in the second period, we now analyze her potential behavior in the first period. There, we distinguish between the two cases $\tilde{\alpha}_2 < \bar{\alpha}_2$ and $\tilde{\alpha}_2 \geq \bar{\alpha}_2$. However, note that these values might differ for the different cases because they rely on the agents' (perceived) search which itself is a function of (perceived) equilibrium outcomes.

a) t=1, $\tilde{\alpha}_2 < \bar{\alpha}_2$

If $\tilde{\alpha}_2 < \bar{\alpha}_2$, the naive agent would expect to be offered $w_2 = w_2^S$ in case both types accept the first-period offer, which we now assume. Hence, the naive type believes to benefit from asymmetric information because his perceived second-period wage is higher than with symmetric information, letting him accept a lower first-period wage:

The naive type's perceived first-period utility then equals

$$w_1 - \frac{1}{2}(s_1)^2 + \beta\delta \left\{ s_1 B + (1-s_1) \left[w_2^S - \frac{1}{2}(\tilde{s}_2^N)^2 + \delta \tilde{s}_2^N B \right] \right\}.$$

Taking into account $\tilde{s}_2^N = \delta B$ and $w_2^S = -\frac{1}{2}(\beta\delta B)^2$, this becomes

$$w_1 - \frac{1}{2}(s_1)^2 + \beta\delta \left\{ s_1 B + (1-s_1)\frac{1}{2}(\delta B)^2 \left(1-\beta^2\right) \right\}.$$

Hence, the naive type's first-period search effort equals

$$s_1 = \beta\delta B \left[1 - \frac{1}{2}\delta^2 B \left(1-\beta^2\right) \right],$$

and the wage giving him a perceived utility of zero amounts to

$$w_1 = -\frac{1}{2}\beta\delta^2 B^2 \left\{ \beta \left[1 - \frac{1}{2}\delta^2 B \left(1-\beta^2\right) \right]^2 + \delta \left(1-\beta^2\right) \right\}. \tag{1}$$

However, the sophisticated agent demands at least the same wage as in the full information case, w_1^S, which is higher than the w_1 just derived.

Thus, for this to be an equilibrium, the principal has to offer w_1^S, the sophisticated agent's full information wage. In terms of the original likelihood of facing a naive agent, α_1, the condition $\tilde{\alpha}_2 < \bar{\alpha}_2$ becomes

$$\begin{aligned}\tilde{\alpha}_2 =& \frac{\alpha_1(1-s_1^N(w_2^S))}{\alpha_1(1-s_1^N(w_2^S))+(1-\alpha_1)(1-s_1^S(w_2^S))} < \bar{\alpha}_2 \\ \Leftrightarrow \alpha_1 <& \frac{\bar{\alpha}_2(1-s_1^S(w_2^S))}{(1-s_1^N(w_2^S))-\bar{\alpha}_2(s_1^S(w_2^S)-s_1^N(w_2^S))} \equiv \bar{\alpha}_1^S\end{aligned}$$

Note that, since $s_1^N(w_2^S) < s_1^S(w_2^S)$, $\bar{\alpha}_1^S < \bar{\alpha}_2$.

Hence, for $\alpha_1 < \bar{\alpha}_1^S$, both types are paid and accept the sophisticated agent's symmetric info wages. The sophisticated type also searches the same amount as with symmetric information, the naive type searches less than the sophisticated agent in the first period. In this case, the principal is strictly better off than with symmetric information. The reason is that her profits when facing the sophisticated type are the same; her profits when facing the naive type are even higher than with the sophisticated type (because of the lower first-period search effort), whereas with symmetric information it is the other way around.

Finally, offering w_1^S is indeed optimal for the principal if $\tilde{\alpha}_2 < \bar{\alpha}_2$. The reason is that a lower wage would not be accepted by the sophisticated agent (and the naive type would be aware of that). Then, the naive type would perceive the principal's second period likelihood of facing himself (conditional on not having received an outside offer) to be equal to 1, in which case he would expect to be offered $\tilde{w}_2^N$. Thus, such a lower wage would also not be accepted by the naive type.

b) t=1, $\tilde{\alpha}_2 \geq \bar{\alpha}_2$

If $\tilde{\alpha}_2 \geq \bar{\alpha}_2$ the naive agent expects the principal to offer $w_2 = \tilde{w}_2^N$ in case he expects both types to go for the first-period wage. Then, the naive type's perceived first-period utility is the same as with symmetric information. Thus, his first-period search effort is also the same as with symmetric information, s_1^N, and he would not accept a first-period offer below the symmetric-info wage, w_1^N.

In this case, the principal has two choices. Either, she offers w_1^N, which however has to be paid to the sophisticated type as well. Or, she sticks to offering the lower wage that keeps the sophisticated agent to his outside option, irrespective of what the naive type might do. We first derive conditions for the first option to be an equilibrium, and then discuss the second option.

Thus, assume that offering w_1^N indeed is an equilibrium when $\tilde{\alpha}_2 \geq \bar{\alpha}_2$, hence the naive agent expects a second-period wage $\tilde{w}_2^N$. In this case, the naive type also expects the sophisticated agent to reject the second-period offer and adjust his search level accordingly. More precisely, he thinks that the sophisticated type – just like himself – has a zero-continuation utility from today's perspective, thus also chooses a search level $\tilde{s}_1^S(\tilde{w}_2^N) = s_1^N(\tilde{w}_2^N) = \beta\delta B$ and also has a first-period reservation wage $w_1^N = -\frac{1}{2}(\beta\delta B)^2$. Therefore, the naive agent perceives the principal's second-period probability of facing a naive type to be

$$\tilde{\alpha}_2 = \frac{\alpha_1(1-\tilde{s}_1^N(\tilde{w}_2^N))}{\alpha_1(1-s_1^N(\tilde{w}_2^N)) + (1-\alpha_1)(1-\tilde{s}_1^N(\tilde{w}_2^N))} = \alpha_1,$$

and the condition $\tilde{\alpha}_2 \geq \bar{\alpha}_2$ becomes $\alpha_1 \geq \bar{\alpha}_2$. Note that, since the sophisticated agent is actually searching less, the principal's true probability of facing a naive agent will be smaller than α_1.

Now, a deviation from offering w_1^N must not be optimal for the principal. Analyzing the respective condition, we must compute the principal's on- and off-path profits. Here, those amount to "real" profits, hence incorporate both type's real behavior. For off-path profits, it is also important to specify the naive agent's belief on the principal's second-period behavior given she deviates.

Offering $w_1^N = -\frac{1}{2}(\beta\delta B)^2$ and hiring both types yields profits

$$\begin{aligned}\Pi_1 =& \pi - w_1^N + \alpha_1\delta\left\{s_1^N\underline{\pi}(1+\delta) + (1-s_1^N)\left[\pi - w_2^N + \delta\left(s_2^N\underline{\pi} + (1-s_2^N)\pi\right)\right]\right\} \\ &+ (1-\alpha_1)\delta\left\{s_1^S\underline{\pi}(1+\delta) + (1-s_1^S)\left[\pi - w_2^S + \delta\left(s_2^S\underline{\pi} + (1-s_2^S)\pi\right)\right]\right\}.\end{aligned}$$

If the principal instead offers $w_1^S < w_1^N$, this can only be an equilibrium if the naive type does not accept the contract (otherwise, offering w_1^S will in *any* case be optimal for the principal). The naive agent will indeed refrain from doing so if he believes that the principal then also offers $\tilde{w}_2^N$ in the second period, in which case the principal's profits are

$$\begin{aligned}\Pi_1 =& \alpha_1\underline{\pi}\left(1+\delta+\delta^2\right) \\ &+ (1-\alpha_1)\left\{\pi - w_1^S + \delta s_1^S\underline{\pi}(1+\delta) + \delta(1-s_1^S)\left[\pi - w_2^S + \delta\left(s_2^S\underline{\pi} + (1-s_2^S)\pi\right)\right]\right\}.\end{aligned}$$

This implies that the principal offers w_1^N if

$$\alpha_1 \geq \frac{w_1^N - w_1^S}{\pi - \underline{\pi} - w_1^S + \delta(1-s_1^N)\left[\pi - \underline{\pi} - w_2^N + \delta(1-s_2^N)(\pi - \underline{\pi})\right]} \equiv \bar{\alpha}_1^P,$$

and if he expects the naive agent to reject any offer below w_1^N. The latter would be supported by the naive type believing that the principal in any case offers $\tilde{w}_2^N$ in the second period.

Now, let us assume that $\alpha_1 < \bar{\alpha}_1^P$. In this case, it is a dominant strategy for the principal to offer w_1^S. The reason is that this yields higher profits than w_1^N even if the naive type rejects the offer. In case the naive type accepts it with a positive probability, her profits would be even larger.

Note that both cases, $\bar{\alpha}_1^S < \bar{\alpha}_1^P$ or $\bar{\alpha}_1^S \geq \bar{\alpha}_1^P$ are possible. Thus, the following cases are feasible:

- $\alpha_1 < \bar{\alpha}_1^P$: The principal offers w_1^S. If $\alpha_1 \leq \bar{\alpha}_1^S$, the naive agent accepts this contract in the first period because he expects the principal to offer w_2^S in the second period. In this case, the principal's profits are higher with asymmetric information than with symmetric information. If $\alpha_1 > \bar{\alpha}_1^S$, the naive agent's decision depends on his beliefs about the principal's second-period offer. If he expects the principal to offer w_2^S in the second period, he accepts. Otherwise, he declines. In the latter case, the principal's profits are smaller than with symmetric information because the naive agent is not employed (whereas the search effort and wage of the sophisticated type are as with symmetric information). Here, we cannot say more on which equilibrium to expect without imposing additional assumptions on the naive agent's belief regarding the principal's belief. The reason is the following circular argument: If the naive agent does not accept the principal's offer and thinks that the principal is aware of that, he expects the principal to offer w_2^S in the second period (because only the sophisticated type would be left). But then, it would be optimal for the naive agent to also accept w_1^S; in this case, however, it would again become optimal for the principal (from the naive agent's view) to offer $\tilde{w}_2^N$ in the second period, and for the naive agent to decline, and so on. Also note that equilibria in mixed strategies would not deliver clearer results: Assume the naive agent expects the principal to mix between w_2^S and $\tilde{w}_2^N$, to make him indifferent between accept and reject in period 1. Then, the first-period wage must be adjusted to convince the naive agent that it will also be accepted by the sophisticated agent (otherwise, he would expect a second-period wage $\tilde{w}_2^N$ with probability 1). Due to the latter's adjusted search effort (again from the naive's perspective), though, the respective first-period wage would have to be so high that the naive agent would strictly prefer to accept it for *any* mixing probability.

- $\alpha_1 \geq \bar{\alpha}_1^P$: If $\alpha_1 \leq \bar{\alpha}_1^S$, the principal offers w_1^S, which is also accepted by the naive type. Then, the principal's profits are larger with asymmetric information. If $\alpha_1 > \bar{\alpha}_1^S$, the principal offers w_1^N if she expects the naive type to reject a lower wage. In this case, the principal's profits are smaller than with symmetric information. The reason is that w_1^N is paid to both types, and search effort and realized outcomes from the second period on are as with symmetric information. With symmetric information, however, the sophisticated agent would be paid the lower w_1^S.

Separating Equilibrium Finally, we sketch a potential separating equilibrium with both types accepting a first-period offer, but also argue that the principal would always prefer a pooling equilibrium if the conditions for the separating equilibrium are satisfied.

Suppose the principal offers two different wages in the first period, one high and one low wage (w_1^H and w_1^L). Then, the sophisticated will always choose the higher wage, since he knows that the second-period wage is independent of his choice. Thus, a separating equilibrium can only exist if the naive agent accepts the low first-period wage.

The naive agent, however, generally has an incentive to imitate the sophisticated agent because he perceives the latter to receive a higher second-period wage in the second period. This implies that he would also go for w_1^H in case he expects the sophisticated type to do so. Therefore, a separating equilibrium can only be sustained if the naive agent expects the sophisticated agent to select w_1^L and consequently does so as well. Such a separating equilibrium, with both types accepting a first-period offer, would hence involve the naive type believing that a pooling equilibrium with both types choosing w_1^L is played, whereas types are effectively separated.

First, the naive type is only willing to select w_1^L if he expects to receive w_2^S in the second period. Thus, $\tilde{\alpha}_2 < \bar{\alpha}_2$ (as derived above) must hold, which we assume from now on (otherwise, also the naive agent would always select the higher wage in the first period, irrespective of his beliefs regarding the sophisticated type's behavior). Second, we assume that the naive agent perceives the principal to assign probability 1 to facing the naive type in case an agent selects w_1^H (otherwise, incentives to choose w_1^L would be weakened) and consequently offer $\tilde{w}_2^N$ in the second period. We can do so in a Bayesian equilibrium because the choice of w_1^H is an off-path event from the naive agent's perspective.

Now, we have shown before that, if the naive agent expects to receive w_2^S in the second period, his first-period search effort equals $\beta\delta B\left[1-\frac{1}{2}\delta^2 B\left(1-\beta^2\right)\right]$. Thus, his perceived utility when choosing w_1^L amounts to

$$w_1^L + \frac{1}{2}\beta\delta^2 B^2 \left\{\beta\left[1-\frac{1}{2}\delta^2 B\left(1-\beta^2\right)\right]^2 + \delta\left(1-\beta^2\right)\right\}. \tag{2}$$

In contrast, if he expects to be paid w_2^N in the second period, his first-period search effort equals $\beta\delta B$. Thus, his perceived utility when choosing w_1^H amounts to

$$w_1^H + \frac{1}{2}(\beta\delta B)^2.$$

Consequently, he is only willing to accept w_1^L if

$$w_1^H - w_1^L \leq \frac{1}{2}\beta\delta^2 B^2 \left\{\delta\left(1-\beta^2\right)(1-\beta\delta B) + \frac{1}{4}\beta\delta^4 B^2\left(1-\beta^2\right)^2\right\} \tag{3}$$

holds. Moreover, the naive agent expects the sophisticated agent to also choose the low wage if going for w_1^H (and consequently rejecting $\tilde{w}_2^N$ in the second period because $\tilde{w}_2^N < w_2^S$) is not optimal for him. The sophisticated type's search effort when expecting to receive w_2^S in the second period amounts to $\beta\delta B\left[1-\beta\delta^2 B(1-\beta)\right]$. Thus, his utility when choosing w_1^L equals (from the perspective of the naive agent)

$$w_1^L + \frac{1}{2}(\beta\delta B)^2\left[1-\beta\delta^2 B(1-\beta)\right]^2 + \delta(1-\beta)\beta^2\delta^2 B^2. \tag{4}$$

If the sophisticated agent selects w_1^H, the naive type expects him to choose a first-period search effort of $\beta\delta B$ and also receive a utility of

$$w_1^H + \frac{1}{2}(\beta\delta B)^2.$$

Thus, the naive agent expects the sophisticated agent to choose w_1^L if

$$w_1^H - w_1^L \leq \frac{1}{2}(\beta\delta B)^2\left[2\delta(1-\beta)(1-\beta\delta B) + \beta^2\delta^4 B^2(1-\beta)^2\right] \tag{5}$$

holds. Now, the sophisticated agent's utility after receiving w_1^L is (from the naive's perspective) smaller than the naive agent's utility in that case. Thus, the smallest feasible w_1^L for the naive agent to believe that the sophisticated type will choose it is determined by setting condition (4) equal to zero. Then, any w_1^H satisfying (3) and (5) could be chosen. Such a w_1^H exists because the right hand sides of both conditions are positive. Indeed, the principal could set $w_1^H = w_1^L + \varepsilon$, with $\varepsilon > 0$ but sufficiently small. In this case, the naive type would accept w_1^L and the sophisticated type w_1^H.

However, note that the w_1^L derived by setting (4) equal to zero is the same as w_1^S, the sophisticated agent's first-period wage under symmetric information, and consequently also the same as the wage given by condition (1), which is paid in a pooling equilibrium in which $\tilde{\alpha}_2 < \bar{\alpha}_2$. Hence, principal would prefer a pooling equilibrium for any $\varepsilon > 0$. ■

A.8 Proof of Proposition 5

First, recall that without a minimum wage,

$$
\begin{aligned}
\tilde{w}_2^N &= -\frac{1}{2}(\delta B)^2 \\
< w_1^N &= w_2^N = w_2^S = -\frac{1}{2}(\beta\delta B)^2 \\
< w_3^S &= w_3^N = 0.
\end{aligned}
$$

Moreover,

$$s_1^N = \beta\delta\left(B - \tilde{w}_2^N - \frac{1}{2}(\delta B)^2\right).$$

From now on, assume $\overline{w} \geq -\frac{1}{2}(\delta B)^2$, hence $\tilde{w}_2^N = \overline{w}$ and

$$\frac{ds_1^N}{d\overline{w}} < 0.$$

Furthermore, recall that the naive agent's perceived period-1 utility equals

$$\tilde{U}_1^N = w_1^N - \frac{1}{2}(s_1^N)^2 + \beta\delta\left\{s_1^N B + (1 - s_1^N)\left[\overline{w} + \frac{1}{2}(\delta B)^2\right]\right\},$$

and w_1^N is the lowest wage that satisfies $\tilde{U}_1^N \geq 0$, or

$$w_1 = \max\left\{\overline{w}; -\frac{1}{2}(\beta\delta)^2\left(B - \overline{w} - \frac{1}{2}(\delta B)^2\right)^2 - \beta\delta\left[\overline{w} + \frac{1}{2}(\delta B)^2\right]\right\}.$$

There, note that
$-\frac{1}{2}(\beta\delta)^2\left(B - \overline{w} - \frac{1}{2}(\delta B)^2\right)^2 - \beta\delta\left[\overline{w} + \frac{1}{2}(\delta B)^2\right]$
is decreasing in $\overline{w}$ for $\overline{w} \geq -\frac{1}{2}(\delta B)^2$, and equals $-\frac{1}{2}(\beta\delta B)^2$ for $\overline{w} = -\frac{1}{2}(\delta B)^2$. Thus, a threshold $\hat{\overline{w}} > -\frac{1}{2}(\delta B)^2$ exists such that $w_1 > \overline{w}$ for $\overline{w} < \hat{\overline{w}}$.

For later use, also note that $\hat{\overline{w}}$ might be larger or smaller than $w_2^{N/S} = -\frac{1}{2}(\beta\delta B)^2$, depending on parameter values.

Now assume $\overline{w} < \hat{\overline{w}}$, hence

$$w_1^N = -\frac{1}{2}(\beta\delta)^2\left(B - \overline{w} - \frac{1}{2}(\delta B)^2\right)^2 - \beta\delta\left[\overline{w} + \frac{1}{2}(\delta B)^2\right]$$

and

$$\frac{dw_1^N}{d\overline{w}} = -\beta\delta\left(1-\beta\delta\left[B-\left(\overline{w}+\frac{1}{2}(\delta B)^2\right)\right]\right) < 0,$$

where the inequality follows from
$\left[B-\left(\overline{w}+\frac{1}{2}(\delta B)^2\right)\right] \leq \left[B-\left(-\frac{1}{2}(\delta B)^2+\frac{1}{2}(\delta B)^2\right)\right] = B.$

Moreover, the naive agent's realized utility equals

$$U_1^N = w_1^N - \frac{1}{2}(s_1^N)^2 + \beta\delta\left\{s_1^N B + (1-s_1^N)\left[w_2^N + \frac{1}{2}(\delta B)^2\right]\right\}.$$

To assess the effect of $\overline{w}$ on U_1^N, we explore the following cases:

- **Case 1:** Either $\overline{w} \leq \hat{\overline{w}} \leq -\frac{1}{2}(\beta\delta B)^2$, or $\hat{\overline{w}} > -\frac{1}{2}(\beta\delta B)^2$ and $\overline{w} \leq -\frac{1}{2}(\beta\delta B)^2$. Then, $w_2^N = -\frac{1}{2}(\beta\delta B)^2 > \tilde{w}_2^N = \overline{w}$.
- **Case 2:** $\hat{\overline{w}} > -\frac{1}{2}(\beta\delta B)^2$ and $\overline{w} \in \left(-\frac{1}{2}(\beta\delta B)^2, \hat{\overline{w}}\right]$. Then, $w_2^N = \tilde{w}_2^N = \overline{w}$.
- **Case 3:** $\overline{w} > \max\left\{-\frac{1}{2}(\beta\delta B)^2, \hat{\overline{w}}\right\}$. Then, $w_1^N = w_2^N = \tilde{w}_2^N = \overline{w}$

Case 1 Now, $ds_1^N/d\overline{w} = -\beta\delta$ yields

$$\begin{aligned}
\frac{dU_1^N}{d\overline{w}} &= \frac{dw_1^N}{d\overline{w}} - s_1^N\frac{ds_1^N}{d\overline{w}} + \beta\delta\frac{ds_1^N}{d\overline{w}}\left[B-\beta\delta^2B^2(1-\beta)\right] \\
&= -\beta^2\delta^2\left(\overline{w}+\frac{1}{2}(\delta B)^2\right) - \beta\delta\left(1-s_1^N\right) + \beta^3\delta^4B^2(1-\beta) \\
&= -\beta\delta\left\{1-2\beta\delta\left[B-\left(\overline{w}+\frac{1}{2}(\delta B)^2\right)\right] + \beta\delta\left[B-\beta\delta^2B^2(1-\beta)\right]\right\} \\
&= -\beta\delta\left\{1-\beta\delta B+2\beta\delta\overline{w}+\beta\delta(\delta B)^2(1-\beta(1-\beta))\right\} < 0,
\end{aligned}$$

where the last inequality follows from $\overline{w} \geq -\frac{1}{2}(\delta B)^2$.

Case 2 Now,

$$\begin{aligned} U_1^N =& w_1^N - \frac{1}{2}(s_1^N)^2 + \beta\delta \left\{ s_1^N B + (1 - s_1^N) \left[\overline{w} + \frac{1}{2}(\delta B)^2 \right] \right\} \\ =& - \left[\beta\delta \left(B - \overline{w} - \frac{1}{2}(\delta B)^2 \right) \right]^2 - \beta\delta \left[\overline{w} + \frac{1}{2}(\delta B)^2 \right] \\ & + \left[\beta\delta \left(B - \overline{w} - \frac{1}{2}(\delta B)^2 \right) \right]^2 + \beta\delta \left[\overline{w} + \frac{1}{2}(\delta B)^2 \right] \\ =& 0, \end{aligned}$$

and

$$\frac{dU_1^N}{d\overline{w}} = 0.$$

Case 3 Now,

$$\begin{aligned} U_1^N =& \overline{w} + \frac{1}{2}(s_1^N)^2 + \beta\delta \left[\overline{w} + \frac{1}{2}(\delta B)^2 \right] \\ =& \overline{w} + \frac{1}{2} (\beta\delta)^2 \left(B - \overline{w} - \frac{1}{2}(\delta B)^2 \right)^2 + \beta\delta \left(\overline{w} + \frac{1}{2}(\delta B)^2 \right) \end{aligned}$$

and

$$\begin{aligned} \frac{dU_1^N}{d\overline{w}} =& 1 + \beta\delta - (\beta\delta)^2 \left(B - \overline{w} - \frac{1}{2}(\delta B)^2 \right) \\ >& 1 + \beta\delta - (\beta\delta)^2 \left(B + \frac{1}{2}(\beta\delta B)^2 - \frac{1}{2}(\delta B)^2 \right) \\ =& 1 + \beta\delta \left[1 - \beta\delta \left(B - \frac{1}{2}(\delta B)^2 \left(1 - \beta^2\right) \right) \right] > 0. \end{aligned}$$

The results for the sophisticated agent follow: He is only affected by a binding minimum wage; if the minimum wage exceeds w_1^S (his first-period wage without a minimum wage), U_1^S goes up. If in addition it exceeds w_2^S, also his search effort is reduced. ■

A.9 Search Benefit

We now relax some assumptions regarding the benefits of search and show that our results continue to hold. First, the agent may continue to search even after having received an offer in a previous period. Second, similar to DellaVigna (2009), successful search allows the agent to draw from an exogenous offer distribution $F(b)$

(with support $[0, \bar{b}]$, where $\bar{b} > 0$). Such an offer yields a minimum payoff b the agent must receive in the next and all following periods (thus, we implicitly also relax the assumption that the total benefit from search is the same in every period). For example, b might reflect the agent's improved position on the labor market and consequently be equivalent to a permanent increase of his outside option. In the following, we will thus refer to b as the agent's outside option. Here, we also allow for successful counteroffers, depending on whether b is smaller or larger than $\pi - \underline{\pi}$. For the following analyse, we do not further pursue this issue; a successful counteroffer increases the agent's net utility from being employed by the same amount, no matter whether he stays with the principal or moves to another employer.

Now, let us denote the average offer as $B \equiv \mathbb{E}(b) = \int_0^{\bar{b}} bdF(b)$. We first consider the sophisticated agent. In period 3, as before, the agent receives a wage which exactly matches his outside option, since there is no further search. In stage 2, conditional on having a per-period outside option b, the agent's search level maximizes

$$-\frac{1}{2}s_2^2 + \beta\delta \left[s_2 \left(\int_b^{\bar{b}} tdF(t) + bF(b) \right) + (1 - s_2)b \right].$$

Thus, the sophisticated agent's optimal period-2 search $s_2^S(b)$ is

$$\begin{aligned} s_2^S(b) &= \beta\delta \left(\int_b^{\bar{b}} tdF(t) - b(1 - F(b)) \right) \\ &= \beta\delta \int_b^{\bar{b}} (t - b)dF(t). \end{aligned}$$

The agent only accepts an offer in period 2 if the following condition is statisfied:

$$\begin{aligned} w_2^S - \frac{1}{2}s_2^S(b)^2 + \beta\delta \left[s_2^S(b) \left(\int_b^{\bar{b}} tdF(t) + bF(b) \right) + (1 - s_2(b))b \right] \geq b + \beta\delta b \\ \Rightarrow w_2^S(b) = b + \frac{1}{2}s_2^S(b)^2 - \beta\delta s_2^S(b) \int_b^{\bar{b}} (t - b)dF(t). \end{aligned}$$

In period 1, the sophisticated agent correctly anticipates his future present bias. As in our main analysis, the expected benefit from search evaluated from the perspective of period 1 is larger than from the perspective of period 2. This also implies that is period-2 utility from the perspective of period 1, $\hat{U}_2^S$, is larger than the the present value of the expected benefits from search,

$$\hat{U}_2^S = B(1+\delta) + \beta(1-\beta)\delta^2 \mathbb{E}\left[(\int_b^{\bar{b}} (t-b)dF(t))^2 \right] > B(1+\delta).$$

Therefore, optimal search of the sophisticated agent in period 1 is:

$$s_1^S = \beta\delta B(1+\delta) - \beta(1-\beta)\delta^2 \left[s_2(0)B - \mathbb{E}\left[s_2(b) \int_b^{\bar{b}} (t-b)dF(t) \right] \right] < \beta\delta B(1+\delta)$$

Moreover, the principal offers him the following wage:

$$\begin{aligned} w_1^S =& \frac{1}{2}(s_1^S)^2 - \beta\delta(s_1^S B(1+\delta) + s_1^S \beta(1-\beta)\delta^2 \mathbb{E}\left[(\int_b^{\bar{b}} (t-b)dF(t))^2 \right] \\ &+ (1-s_1^S)\beta(1-\beta)\delta^2 B^2)) \\ <& \frac{1}{2}(s_1^S)^2 - \beta\delta s_1^S B(1+\delta) \end{aligned}$$

Summarizing, the principal can push the period-1 wage of the sophisticated agent below the (negative) net surplus of search not only due to the "extra" utility of assessing period-2 search from the perspective of period 1 for the case that search has not been successful, but also for the case of receiving a better offer. Thus, his first-period utility continues to be $U_1^S = 0$.

We now turn to the analysis of the naive agent. As in our main analysis, his realized second- (and third-)period outcomes are as for the sophisticated agent. Still, he believes to be time-consistent in future periods and thus expects his period-2 employer to push him to his perceived outside option. This implies that his expected search and wage level conditional on having received an offer b are:

$$\begin{aligned} \tilde{s}_2^N(b) &= \delta \int_b^{\bar{b}} (t-b)dF(t) \\ \tilde{w}_2^N(b) &= b + \frac{1}{2}\tilde{s}_2(b)^2 - \delta\tilde{s}_2(b) \int_b^{\bar{b}} (t-b)dF(t) \end{aligned}$$

Consequently, his wage and search level in period 1 are:

$$\begin{aligned} s_1^N &= \beta\delta(1+\delta)B \\ w_1^N &= \frac{1}{2}(s_1^N)^2 - \beta\delta s_1^N(1+\delta)B \end{aligned}$$

As in our main model, the naive agent not only searches more than the sophisticated agent, but also has a positive realized utility,

$$U_1^N = \beta\delta\left(s_1^N \beta(1-\beta)\delta^2 \mathbb{E}\left[(\int_b^{\bar{b}} (t-b)dF(t))^2\right] + (1-s_1^N)\beta(1-\beta)\delta^2 B^2\right) > 0$$

■

B Supplementary Appendix

B.1 Partially Naive Agent

Following Eliaz and Spiegler (2006), we model partial naivete as *frequency naivete*: with probability $\theta \in [0,1]$, the period-1 agent perceives his period-2 self to *not* be present biased; with the remaining probability $1-\theta$, the period-1 agent correctly perceives his period-2 self to additionally discount future payoffs with β. $\theta = 1$ describes a fully naive agent who thinks that his present bias disappears in the next period with probability 1, and that he discounts the future exponentially from then on. In contrast, $\theta = 0$ describes a sophisticated agent who perfectly anticipates his future present bias. Here, we show that a lower θ, i.e., more sophistication, monotonously reduces search as well as the agent's realized utility level.

Outcomes in periods two and three are independent of θ, hence the same as with a sophisticated and fully naive agent. From the perspective of period 1, a partially naive agent expects to maximize $-(s_2)^2/2+\delta s_2 B$ in period $t=2$ with probability θ, and $-(s_2)^2/2+\beta\delta s_2 B$ with probability $1-\theta$. This implies that the partially naive agent expects to choose a search level $\tilde{s}_2(\theta)$ which is characterized by

$$\tilde{s}_2 = \delta B \text{ with probability } \theta$$
$$\tilde{s}_2 = \beta\delta B \text{ with probability } 1-\theta.$$

Furthermore, in period 1 the partially naive agent expects to be offered a second-period wage

$$\tilde{w}_2^{PN} = -\frac{1}{2}(\delta B)^2 \text{ with probability } \theta$$
$$\tilde{w}_2^{PN} = -\frac{1}{2}(\beta\delta B)^2 \text{ with probability } 1-\theta.$$

The agent's behavior in $t=1$ is determined by his perceptions of future outcomes, not their true realizations.

This yields

Lemma 4 *A partially naive agent with $\theta \in (0,1)$*

- *exerts less search effort in the first period than in the second, i.e. $s_1^{PN} < s_2^{PN}$.*
- *receives a first-period wage that is* lower *than the first-period negative search benefit, i.e.$w_1^{PN} < \frac{1}{2}(s_1^{PN})^2 - \beta\delta s_1^{PN} B$.*

Moreover, in the first period a higher extent of naivete lets an agent search more and receive a higher wage, i.e. $ds_1^{PN}/d\theta > 0$ and $dw_1^{PN}/d\theta > 0$.

Proof. A partially naive agent perceives his first-period utility to be (already taking into account $w_3 = 0$)

$$\tilde{U}_1^{PN} = w_1 - \frac{1}{2}(s_1)^2 + \beta\delta \left\{ s_1 B + (1 - s_1) \left[\theta \cdot 0 + (1 - \theta)\beta\delta^2 B^2 (1 - \beta) \right] \right\}.$$

The first-order condition yields

$$s_1^{PN} = \beta\delta \left[B - (1 - \theta)\beta\delta^2 B^2 (1 - \beta) \right] < s_2^{PN} = \beta\delta B,$$

with

$$\frac{ds_1^{PN}}{d\theta} = \beta^2\delta^3 B^2 (1 - \beta) > 0.$$

Moreover,

$$\begin{aligned} w_1^{PN} &= \frac{1}{2}(s_1^{PN})^2 - \beta\delta \left[s_1^{PN} B + (1 - s_1^{PN})(1 - \theta)\beta\delta^2 B^2 (1 - \beta) \right] \\ &< \frac{1}{2}(s_1^{PN})^2 - \beta\delta s_1^{PN} B \end{aligned}$$

and

$$\frac{dw_1^{PN}}{d\theta} = \beta^2\delta^3 B^2 (1 - \beta) \left(1 - \beta\delta B + (1 - \theta)\beta^2\delta^3 B^2 (1 - \beta) \right) > 0.$$

■

Finally, Proposition 6 explores the effect of an agent's naivete on his realized utility.

Proposition 6 *The utility of a partially naive agent is positive and strictly increasing in θ.*

Proof. A partially naive agent realized period-1 utility equals

$$\begin{aligned} U_1^{PN} =& w_1^{PN} - \frac{1}{2}(s_1^{PN})^2 \\ &+ \beta\delta \left\{ s_1^{PN} B + (1 - s_1^{PN}) \left[w_2^{PN} - \frac{1}{2}(s_2^{PN})^2 + \delta \left(s_2^{PN} B + (1 - s_2^{PN}) w_3^{PN} \right) \right] \right\}. \end{aligned}$$

Taking into account $w_1^{PN} = \frac{1}{2}(s_1^{PN})^2 - \beta\delta \left[s_1^{PN} B + (1 - s_1^{PN})(1 - \theta)\beta\delta^2 B^2 (1 - \beta) \right]$, $w_2^{PN} = \frac{1}{2}(s_2^{PN})^2 - \beta\delta s_2^{PN} B$, $s_2^N = \beta\delta B$ and $w_3^{PN} = 0$,

$$U_1^{PN} = \theta(1 - s_1^{PN})\beta^2\delta^3 B^2 (1 - \beta),$$

with

$$\begin{aligned}
\frac{dU_1^{PN}}{d\theta} &= (1-s_1^{PN})\beta^2\delta^3B^2(1-\beta) - \frac{ds_1^{PN}}{d\theta}\theta\beta^2\delta^3B^2(1-\beta) \\
&= \beta^2\delta^3B^2(1-\beta)\left[1-\beta\delta B+(1-2\theta)\beta^2\delta^3B(1-\beta)\right] \\
&\geq \beta^2\delta^3B^2(1-\beta)\left[1-\beta\delta B-\beta^2\delta^3B(1-\beta)\right] \\
&= \beta^2\delta^3B^2(1-\beta)\left[1-\beta\delta B\left(1+\beta\delta^2(1-\beta)\right)\right] \\
&\geq \beta^2\delta^3B^2(1-\beta)\left[1-\delta B\beta\left(1+\beta(1-\beta)\right)\right] \\
&= \beta^2\delta^3B^2(1-\beta)\left[1-\delta B\left(1-(1-\beta)\left(1-\beta^2\right)\right)\right] \\
&> 0,
\end{aligned}$$

where the latter follows from $\delta B \leq 1$.

■

Proposition 4 indicates that our results on the differences between a sophisticated and a fully naive agent hold monotonically, for any value $\theta \in (0,1)$.