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

Matthias Fahn† & Regina Seibel‡

December 11, 2019

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.

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, 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. Whereas employees with inconsistent time preferences generally search too little from the perspective of earlier periods, those who are naive about their present bias might actually *search more* and be *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, and consequently stickiness, of the generated inefficiencies (exceptions are DellaVigna and Paserman, 2005; Paserman, 2008).

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

[1]See DellaVigna (2009) or Kőszegi (2014) 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).

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 perceives his search effort to be 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). Whereas the sophisticated agent perfectly anticipates his future present bias, the naive agent expects to be time-consistent later on. 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. This affects the size of feasible wage cuts to exploit the benefits of on-the-job search. The second-period wage completely extracts the agent's expected net benefits from second-period 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 second-period wage. But this is only fully recognized by a sophisticated 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 in period 1 and consequently does not perceive his period-2

[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 5.

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 first-period search effort. Since the first-period 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.

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, in particular if a firm cannot commit to long-term contracts. 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,[7] 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 test in the hiring process which make use of artificial intelligence and allow firms to develop a better idea about their employees' characteristics might also have negative side effects. This result is again driven by the naive agent's misperception of second-period outcomes. First note that a separation of types with both accepting the first-period offer is never optimal for the principal. Thus, if the naive agent expects both types to still be around at 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 second-period rent and thus accept a lower first-period wage than with symmetric information. In this case, the principal's asymmetric-info profits would indeed be larger. If the share of naive

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

[7] Such as Eliaz and Spiegler, 2006; Heidhues and Kőszegi, 2010; Heidhues and Kőszegi, 2017

agents is larger, the second option seems optimal, and the principal will generally be harmed by asymmetric information. Then, she might offer a higher first-period wage, and sophisticated agents benefit from the existence of naive agents.[8] Alternatively, she might make an offer that the naive agent potentially rejects.

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* 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.

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 5). 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,

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

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 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 of 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 5 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.[9] 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.

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

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 first characterize equilibria for sophisticated and fully naive agents separately and subsequently compare the two outcomes.

3.1 Sophisticated Agent

First, 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. 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 $w_3^S = 0$, i.e., his net utility of being employed just equals his outside utility of zero.

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,

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

Consequently, on-the-job search can be regarded as a non-pecuniary benefit of being employed that allows the principal to reduce the second-period wage below the agent's reservation utility (as previously derived by Board and Meyer-Ter-Vehn, 2015). Moreover, the agent's time-inconsistency gives the principal additional, intertemporal, opportunities to reduce wages.

Lemma 1 *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 first-period wage is* lower *than the first-period 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.

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 second-period 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 second-period 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 second-period 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 2 *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 first-period wage is equal to the first-period 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.

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.

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.

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 first-period 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.[10] 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 first-period outcomes.

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

[10] 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.

The proof can be found in Appendix A.

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 refer to as symmetric information.[11] 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 second-period 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 belief 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

[11]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).

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.

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

Now, the results rely on the naive agent's belief about the principal's second-period offer, but also on the principal's belief about the naive's first-period 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 first period 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 second-period search effort. If α_1 is sufficiently small, the former seems optimal. Then, the naive type expects to be paid a higher second-period 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 second period utility of staying with the principal is higher.

If α_1 is higher, the naive agent expects the principal to offer a low second-period wage (in case both types have accepted the first-period contract) and thus exclude the sophisticated type. Then, he is only willing to accept the first-period 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 second period wage but too low for a high first-period 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 first-period 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 first-period 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.[12]

Finally, we could incorporate recent evidence that many humans are naive about their own present bias but correctly predict others' time-inconsistency (Fedyk, 2018). This would make naive agents even more prone to believe that a high second-period wage is optimal for the principal and accept a lower second-period 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 Discussion and Conclusion

We have shown that 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. Moreover, the principal might be better off when being ignorant about the extent of the agent's naiveté. To conclude, we discuss two of our main assumptions, the absence of long-term commitment and the simultaneous assessment of wage and search costs in a given period, and explore the robustness of our results when relaxing them. First, the principal would benefit from being able to commit to future wage payments. Then, she would 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. However, uncertainty regarding future prospects might reduce the benefits of commitment. Assume there is a chance that the principal's profits

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

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, 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 that consumption more than a few days away is not included in the "present" an individual is biased towards. Incorporating these aspects into our model, we might split each period into two steps and discount period-t search with β already *at the beginning* of period t. If the naive agent then also overestimates this period's search, he is willing to accept an extra wage reduction. For example, at the beginning of $t = 2$ he would expect to search more in this period than he actually will and accept a lower wage. His realized utility at the beginning of period 2 will thus be negative. From the perspective of the first period, however, discounting between periods 2 and 3 is lower, which increases the present value of future search benefits and overcompensates for the exploitative period-2 wage. Then, although the principal now can pay the naive agent an exploitative second-period wage, his realized second-period utility from the perspective of period 1 continues to be positive. In this case, the naive agent's realized utility at the beginning of his career would continue to be positive 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.

References

Augenblick, N. (2017). Short-term time discounting of unpleasant tasks. 5

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

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, 3.1

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

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

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, 7, 10, 4, 11, 5

Englmaier, F., Fahn, M., and Schwarz, M. A. (2016). Long-term employment relations when agents are present biased. 2, 10

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

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. 10

Heidhues, P. and Kőszegi, B. (2010). Exploiting naivete about self-control in the credit market. *American Economic Review*, 100(5):2279–2303. 6, 7

Heidhues, P. and Kőszegi, B. (2017). Naïveté-based discrimination. *The Quarterly Journal of Economics*, 132(2):1019–1054. 7

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. 8

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, 10

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

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. 5

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

Appendix A – Omitted Proofs

Proof of Lemma 1. s_1^S maximizes

$$-\frac{1}{2}s_1^2+\beta\delta\left\{s_1B+(1-s_1)\left[w_2-\frac{1}{2}s_2^2+\delta s_2B\right]\right\}$$
$$=-\frac{1}{2}s_1^2+\beta\delta\left\{s_1B+(1-s_1)(1-\beta)\beta\delta^2B^2\right\},$$

hence

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

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

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

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

Proof of Lemma 2 A fully naive agent perceives his first-period utility to be

$$\tilde{U}_1^N=w_1-\frac{1}{2}(s_1)^2+\beta\delta\left\{s_1B+(1-s_1)\left[\tilde{w}_2^N-\frac{1}{2}(\tilde{s}_2^N)^2+\delta\left(\tilde{s}_2^NB+(1-\tilde{s}_2^N)w_3^N\right)\right]\right\}.$$

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

$$\tilde{U}_1^N=w_1-\frac{1}{2}(s_1)^2+\beta\delta s_1B.$$

The Lemma immediately follows. ■

Proof of Proposition 1.

As shown in Lemma 1, 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 2, 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. ■

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 $w_3^N = 0$,

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

■

Proof of Proposition 3. Recall that first-period wages are

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

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

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

Hence

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

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 $\underline{\pi} < \pi$.

■

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,

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

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

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

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

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

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

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

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 $\varepsilon > 0$. ■

Appendix B – 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 3 *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 5 explores the effect of an agent's naivete on his realized utility.

Proposition 5 *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

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

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

■

Proposition 5 indicates that our results on the differences between a sophisticated and a fully naive agent hold monotonically, for any value $\theta \in (0, 1)$.