8 *Underinvestment in training?*

RANDOLPH SLOOF, JOEP SONNEMANS AND HESSEL OOSTERBEEK

8.1 Introduction

When a newly hired employee enters a firm there are typically many skills he has to acquire before becoming fully productive. A blue-collar worker, for example, needs to get acquainted with the machines and tools with which he is going to work, while white-collar workers usually have to become familiar with, for example, the particular software package in place. Many of these skills are highly firm-specific, i.e. they are of much lower value in other firms. Investing in firm-specific skills is therefore a risky enterprise for the worker. Although the employer may promise to reward the worker for skills acquisition, for instance through a promotion or a wage increase, after the investment has been made she has an incentive to renege on this promise in order to save on labour costs. Reneging is possible for the firm, because the worker cannot collect the return to his newly obtained skills elsewhere (and because the labour contract is incomplete: see below). Of course, workers will anticipate this opportunistic behaviour by the firm and are unlikely to invest in the first place. Likewise, the firm has not much of an incentive to make the investment. Once the worker has obtained the skills, he will bargain for a wage increase or a promotion. The firm is in a weak bargaining position, because the specific skills required ensure that the worker cannot immediately be replaced by another one. Part of the returns on investment are thus captured by the worker, reducing the firm's initial incentives to invest.

In this chapter we focus on underinvestment in specific training. By now there is a substantial theoretical literature that analyses underinvestment in firm-specific human capital: see Malcomson (1997, 1999) for overviews. The main focus is on factors that hinder efficient investment in work-related training and on the assessment of potential contractual remedies for these inefficiencies. It has been derived that, in particular circumstances, well-designed contracts and/or other types of

arrangements may protect the specific investments made. Examples include the inclusion of payback clauses in the employment contract – for example, when the worker quits his job he has to reimburse part of the firm's training costs – and particular types of promotion policies like up-or-out: after a probation period a worker is either promoted or dismissed.

Underinvestment is predicted to occur when employers and workers behave selfishly. But in reality this is typically not always the case (cf. Bewley, 1999). The actual extent and importance of the underinvestment problem is therefore ultimately an empirical issue. The same applies as to whether the proposed solutions do indeed work as theory predicts. By the very nature of the issues involved, however, it is difficult to gather field data to investigate this. First, underinvestment results from the fact that specific investments in human capital are not verifiable by a third party like a court. Enforceable agreements about who invests and how the returns and costs are divided therefore cannot be made at the time the worker signs the contract. But if a court cannot verify the acquisition of skills, why should the empirical researcher be able to do so? Indeed, in standard field data research the measurement of the amount of training is highly problematic; see section 3.3 for a clear and concise discussion of the conceptual and practical problems involved.

A second serious limitation of field data is that, even if it were possible to observe the level of investment in skills, it would be extremely hard to determine the actual productivity of a worker and what he would have produced in alternative employment or without training. Yet this information is crucial for assessing both the extent of the underinvestment problem and the solutions to it. As explained in section 3.4, recovering this type of counterfactual information is *the* fundamental problem in estimating the wage returns to training (and thus in establishing underinvestment empirically).

With these kinds of data problems hampering the test of theoretical predictions, laboratory experiments offer an attractive alternative. In experiments almost everything is under the control of the researcher. In particular, the actual investment decision can be observed and the researcher fixes the economic value of the worker's productivity for every possible alternative. Exploiting these advantages, a number of controlled experiments have been run to verify the empirical relevance of the underinvestment problem and the performance of several

contractual solutions to it. Although overall the results obtained are fairly mixed, two general conclusions can be drawn. First, underinvestment is typically much less of a problem than theory predicts it to be. As a result, the efficiency loss due to (less) underinvestment is much smaller than predicted. This finding is in line with the results reported in chapter 3 that the wage returns to training are fairly small.[1] Second, some of the theoretical solutions do indeed alleviate underinvestment, albeit to a lesser extent than predicted. One important reason is that subjects in our experiments behave less opportunistically than standard theory predicts. Other motivational factors like fairness and reciprocity play an important role. Taken together, these results do not indicate that government intervention is called for.

In the next section we discuss the results of one particular set of laboratory experiments in more detail. The set-up and the results of this experiment are representative of other experiments that are conducted within this research area. In the third section we provide a brief review of these other relevant experimental studies. The final section discusses both the limitations and the implications of the experimental results obtained.

8.2 Promotion rules and skill acquisition

A number of organizations make use of so-called *up-or-out* promotion policies. Examples include law firms, partnerships, the military and American universities. Up-or-out requires that, after a probationary period, an employee is either promoted or dismissed. By having such a harsh promotion policy, employees obtain strong incentives to invest in firm-specific skills. This holds because the more skills collected, the more likely it becomes that the employee will make the grade and get promotion. By boosting investment incentives, up-or-out may reduce underinvestment. Unfortunately, this promotion rule is not necessarily a free lunch and may come at potentially high economic costs: it may waste the acquired skills of those not promoted. For example, associates that appear unsuitable for becoming partners may still be very valuable employees for the current law firm; given the firm-specific knowledge they have collected it would result in an inefficient match if they were forced to leave the company and find employment elsewhere.

A well-known textbook on organizational economics formulates the above trade-off as follows (cf. Milgrom and Roberts, 1992: 364):

Promotions serve two roles in an organization. First, they help assign people to the roles where they can best contribute to the organization's performance and success. Second, promotions serve as incentives and rewards. These conceptually distinct roles are sometimes in conflict.

An alternative to up-or-out is to keep those valuable employees that are not promoted within the organization. This results in an *up-or-stay* promotion rule, under which non-promoted workers are offered some lower-level job. In the absence of a strong (and credible) threat of being fired, workers' investment incentives are now muted compared to up-or-out. This especially applies when firms use a no lay-off policy, such as large companies in Japan and US firms like IBM, Hewlett Packard and Lincoln Electric used to have (cf. Milgrom and Roberts, 1992).

The choice between up-or-out and up-or-stay represents the trade-off mentioned above between inefficient matching of workers to jobs and inefficient investments in skills acquisition. To what extent such a trade-off really exists, and whether up-or-out can indeed be used to strengthen investment incentives, is verified experimentally in Oosterbeek et al. (2006). A summary of this research project is given below. For brevity many of the details are left out; the reader interested in an in-depth discussion is referred to the original paper.

The economic environment

To facilitate internal validity and to keep full control, the laboratory experiment itself is very stylized. We first provide an exact description of the situation considered in the lab, before we provide an intuitive interpretation of what kind of real world situations are represented.

In the experiment a worker and an employer interact during two periods. In the first period the worker has the opportunity to make a relationship-specific investment at a cost of 25.[2] This investment increases the probability that he is of high productivity in the second period. Without investment this probability equals 1/4; with investment it increases to 3/4. Before the second period starts the employer has to offer the worker one of two possible job levels: an easy job or a difficult job. Alternatively she may decide to fire the worker. After the investment decision, but before the employer makes her job offer, the worker's actual productivity level is revealed to both parties. When the worker is of low productivity he produces 100 in the easy job and 0 if assigned to a difficult job. When the worker turns out to be of high productivity he

produces 175 if he is assigned to the easy job and 220 in the difficult job. Independent of the productivity level within the firm, the worker's outside productivity equals 0.[3] When the worker is offered either the easy or the difficult job, he can accept or reject this offer. Rejection implies that the worker leaves the firm and obtains his outside wage of 0.

In this situation efficiency requires that low-productivity workers are assigned to the easy job and high-productivity workers to the difficult job. Moreover, making the investment is efficient, because the expected economic benefits of doing so (60) exceed the economic costs of 25. Here the expected benefits can be calculated as follows. The productivity advantage of a high-productivity worker over a low-productivity one equals $220 - 100 = 120$ points. Investment increases the probability of a high-productivity type by 50% (i.e. from 1/4 to 3/4), so on average the benefit of investment equals 50% of 120, i.e. 60.

Investment is efficient, but a priori it is not clear that the worker will indeed invest. Underinvestment may arise because parties can neither contract upon the worker's investment decision nor on his actual productivity. The employer can only attach wage levels to different job levels and can commit to these wages. We assume that the wage in the difficult job equals $w_d = 110$. Hence when the worker is of high productivity and assigned to the difficult job, the overall value of his productivity of 220 is divided equally; the employer gets 110 and the worker gets 110. The wage in the easy job w_e is used as a treatment variable, i.e. it is varied within the experiments, to account for the various types of promotion rules. Table 8.1 summarizes the number of points (gross pay-offs) the two parties obtain in each case, depending on the worker's level of productivity and his job assignment.

Table 8.1 Number of points (gross of investment costs) for employer and worker

	Low-productivity worker		*High-productivity worker*	
Job:	*Employer*	*Worker*	*Employer*	*Worker*
Difficult	−110	110	110	110
Easy	$100 - w_e$	w_e	$175 - w_e$	w_e
Out	0	0	0	0

One real world application the above situation represents is a law firm that recruits associates for junior positions. After a probationary period (period 1 in the experiment) the firm decides whether an associate is suitable for becoming a partner (i.e. is of high productivity). If so, the firm prefers to make him a partner (i.e. offer him the difficult job). If not, i.e. the associate is of low productivity, it would be best to keep him in the firm at the associate's position (i.e. in the easy job). Whether an associate is a suitable candidate for becoming a partner depends on personal characteristics, but also in part on the firm-specific skills obtained during the probationary period. In particular, associates that obtained more skills (i.e. made the investment) have higher chances of being a highly productive partner ($\frac{3}{4} > \frac{1}{4}$). The firm therefore would like to stimulate associates to invest in firm-specific skills. Unfortunately, however, both the associate's investment and his actual productivity are non-verifiable to a court and therefore cannot be part of the employment contract. The only thing the firm can do is to commit itself to certain wage levels belonging to different jobs, i.e. to a particular salary structure. As will be explained below, this de facto comes down to a particular type of promotion rule.

The experimental set-up may capture other situations as well. There actually need not be a big difference between the two different jobs, i.e. the firm may simply generate different 'job titles' for in essence the same type of tasks.[4] Think, for example, of junior and senior researchers, junior and senior managers, civil servants and so on. There is typically not much of a difference between the tasks of a junior employee and a senior employee. These different job titles are mainly generated to justify salary differences, which in turn can be used as an instrument for incentive purposes. Although the set-up in the experiment is highly simplified, it is thus able to capture the essential characteristics of promotion as an incentive device.

The different promotion rules

We return to the specifics of the experimental set-up. Recall that the wage in the difficult job is set equal to $w_d = 110$. The offered wage in the easy job (w_e) then determines the promotion and firing policies the employer will apply. In the experiment we consider three different values of w_e, namely, 110, 70 and 50. As will be explained in detail

below, these three different wage levels correspond to the following three promotion rules:

1. Up-or-out promotion rule: $w_e = 110$ $(= w_d)$.
2. Up-or-stay promotion rule: $w_e = 70$.
3. Stay-or-stay promotion rule: $w_e = 50$.

To understand these promotion rules, assume that both the employer and the worker are only interested in the number of points they get themselves. In period 2 the worker then accepts any job offer, because both the easy and the difficult job yield him more than his outside option of 0. We thus can focus on the employer's job offer. Now consider first the up-or-out rule. With $w_e = 110$ the firm prefers to offer a high-productivity worker the difficult job. This yields the firm more than keeping him in the easy job $(110 > 175 - 110)$ or firing him instead $(110 > 0$, cf. table 8.1). However, when a low-productivity worker is kept, the firm either obtains -110 (difficult job) or $100 - 110 = -10$ (easy job). So the firm is better off firing the low-productivity worker. Workers are thus either promoted or dismissed, explaining the term 'up-or-out'.

The predictions for the other two promotion rules follow similarly. Under up-or-stay the firm is best off assigning the high-productivity worker to the difficult job (because $110 > 175 - 70$) and the low-productivity worker to the easy job. The latter follows because the wage of $w_e = 70$ is now low enough for the firm to make a profit even when the worker is of low productivity. In the case of stay-or-stay $(w_e = 50)$ the worker is always kept in the easy job, irrespective of his productivity. Here the additional wage costs $(110 - 50 = 60)$ of promoting a high-productivity worker are simply too high compared to the productivity increase that results after promotion $(220 - 175 = 45)$.

The promotion rules have different implications for the worker's investment incentives. Under stay-or-stay the worker will never invest, because the costly investment then yields him no private benefits at all. He will always be assigned the easy job and obtain $w_e = 50$. More generally, lacking any promotional prospects, workers have very weak incentives to invest in skills. Under up-or-stay the investment has some private benefits, because it increases the probability of getting the high wage of 110 rather than the low wage of 70 by 50%. Specifically, the worker's expected benefit of investment equals 50% of $(110 - 70)$, i.e. 20. Because this falls short of the investment costs of 25, the worker is still predicted not to invest. In this respect up-or-stay is

Table 8.2 Percentages of efficient decisions

		Up-or-out ($w_e = 110$)	*Up-or-stay* ($w_e = 70$)	*Stay-or-stay* ($w_e = 50$)
Investment:	Predicted	*100%*	*0%*	*0%*
	Actual	*85%*	*46%*[#]	*43%*[#]
Assignment:				
High → Difficult:	Predicted	*100%*	*100%*	*0%*
	Actual	*100%*	*96%*	*69%*
Low → Easy:	Predicted	*0%*	*100%*	*100%*
	Actual	*19%*	*92%*	*97%*

Remark. Within the rows labelled 'Actual' the observed percentages are significantly different (at the 5% level) between the three promotion rules, except for the comparison indicated by superscript # in the second row.

similar to stay-or-stay. The private benefits from investment are highest under the up-or-out promotion rule. The worker's expected return then equals 50% of (110 – 0), i.e. 55, which induces him to invest.

Efficiency requires that the worker makes the investment and that the low- (high-)productivity worker is assigned to the easy (difficult) job. But theory predicts that none of the three promotion rules attains full efficiency.[5] Up-or-out wastes the skills/production of the low-productivity workers that are dismissed rather than kept in the easy job. Under up-or-stay job assignment is efficient, but the worker is predicted to refrain from efficient investment. Stay-or-stay is suboptimal with respect to both investment and assignment. The predicted percentages of efficient decisions are summarized in table 8.2, in the rows labelled 'Predicted'.

Experimental results

The theoretical predictions were tested in the CREED laboratory at the University of Amsterdam. Subjects could only communicate by means of a computer network and did not know with whom they were connected. Most of the 160 subjects that participated in the experiment were undergraduate students in economics (66%). The experiment started with on-line instructions. These instructions, and the

experiment itself, were phrased neutrally; words like employer, worker and investment were avoided. Instead of using these value-laden terms we labelled the employer as the 'participant with role A' and the worker as the 'participant with role B'. The choice whether to invest or not was presented to the subjects as a choice between two different 'disks', with each of them representing visually the corresponding probability distribution of being of low ('blue') or high ('yellow') productivity.

Before the start of the first round all subjects received a message that informed them about their role. Subjects then played the strategic game described earlier 30 times (30 rounds). In particular, for a given promotion rule, workers first decided whether to invest or not. This determined their probability of being of high productivity. The computer then determined the individual worker's actual productivity on the basis of a random draw (using the appropriate probability distribution). Both the worker and the employer were informed about the outcome of this draw. Subsequently employers made their job offers, choosing between the difficult job, the easy job, or firing the worker. In the final stage the worker decided whether to accept the employer's job offer (if any) or not.

In each round subjects were paired with a different opponent. This was done in such a way as to keep the same matchings at a minimum. In particular, by using a rotating matching scheme that best preserved the one-shot nature of the strategic situation, we ruled out reputational considerations. After they had played 30 rounds, subjects filled out a short questionnaire. At the end of the experiment the earned experimental points were exchanged for money at a rate of 1 point = 1 eurocent. Subjects earned on average 22 euros in about one and a half hours.

The main results of the experiment are summarized in the rows labelled 'Actual' in table 8.2. These give, for each promotion rule respectively, the mean propensity to invest, the percentage of high-productivity workers that are efficiently promoted to the difficult job and the percentage of low-productivity workers that are efficiently kept in the easy job. Comparing the actually observed percentages across promotion rules by means of appropriate statistical tests, all differences are significant at the 5% level except for the insignificant difference in investment rates between up-or-stay and stay-or-stay.

We will discuss the outcomes for the three rules in detail below, but two main findings can already be noted from table 8.2. First, the predictions across the three different promotion rules are strongly

supported by the data. In particular, (i) the propensity to invest is significantly higher under up-or-out than under up-or-stay and stay-or-stay (which do not differ in this respect), (ii) inefficient dismissals are significantly more likely under up-or-out and (iii) (efficient) promotions are significantly less likely under stay-or-stay. It thus holds true that up-or-out boosts investment incentives, at the cost of wasting the skills of those not promoted.

A second key finding is that in practice the promotion rules perform differently from what standard theory predicts. This especially holds true for the up-or-stay and stay-or-stay rules, which perform better than predicted. For the parameters chosen, this different performance even changes the efficiency ranking of the three promotion rules. The up-or-stay rule appears to perform best on efficiency grounds (evaluated at total surplus, see below), while standard theory predicts that up-or-out would be optimal. This result can be explained by the fact that the promotion rules differ in the extent to which they give scope to efficiency-enhancing reciprocity. Here reciprocity refers to the motivation to reward fair behaviour and to punish unfair behaviour, even though these rewards and punishments are costly to carry out. Below we discuss this in more detail, by first considering the three different promotion rules in isolation.

Up-or-out ($w_e = 110$)

The outcomes under up-or-out are closest to the theoretical predictions. The investment rate of 85% is fairly close to 100%, all high-productivity workers are assigned to the difficult job and low-productivity workers are typically not retained in the easy job. The assignment of low-productivity workers in fact varies with the actual investment made (this cannot be observed from table 8.2). After no investment 97% are laid off, while after investment only 73% of the low-productivity workers are fired. This assignment behaviour can be explained by positive reciprocity. The firm is sometimes willing to forgo some money in order to reward the worker for his unsuccessful investment. But, given that this channel to reciprocate is very cheap to the firm (it costs the employer 10 points to give a reward of 110 points to the worker), the fact that it is not observed in 73% of the cases suggests that this mechanism is not very strong. The impact in terms of overall efficiency is also rather minor; here efficiency is measured as the overall surplus realized, i.e. the worker's total productivity minus the actual

investment outlays. There is a small efficiency gain because a small fraction of the low-productivity workers are kept rather than dismissed. Yet there is some extra efficiency loss because workers sometimes do not invest. The realized surplus is on average around 135 under the up-or-out contract, while 140 was predicted.

Up-or-stay ($w_e = 70$)

Here the mean propensity to invest equals 46%, while no investments are predicted. The actual assignment patterns can provide an explanation for the higher propensity to invest. Table 8.2 indicates that a very small fraction (4%) of high-productivity workers are not promoted and also that a small fraction (8%) of low-productivity workers are dismissed. What the table does not reveal is that these percentages actually vary with whether the worker made the investment or not. After investment both percentages are very low at 2%, while after no investment they both equal 9%. This assignment pattern can be interpreted as negative reciprocity, in the following way. There are workers who randomly (i.e. with probability ¼) turn out to be of high productivity without making the investment. They are not rewarded for their productivity but rather punished, as they made no effort. In particular, when non-investing workers appear to be of high productivity they are in 9% of the cases offered a less attractive job (easy) than corresponds with their productivity (difficult). The same applies to low-productivity workers who did not invest.

While only a small fraction of the non-investors are actually punished, this fraction is large enough to motivate workers' investment behaviour. Clearly, for this to be the case, workers have to anticipate the actual job offer patterns. Given these actual patterns the expected pay-off from investment equals 74 points, while the expected pay-off from not investing equals about the same. Hence, investment is not irrational at all when the up-or-stay contract applies, because the worker may correctly anticipate the negative reciprocal response of (a small fraction of) employers when he would not invest.

Because actual assignments do not exactly mimic the predicted efficient assignment rule, there are some additional efficiency losses under up-or-stay. The associated losses are, however, small in comparison with the efficiency gain resulting from the unexpectedly high (46%) investment rate. The realized social surplus under up-or-stay on average amounts to 140, where only 130 is predicted under this promotion

rule. Up-or-stay thus performs better than up-or-out in terms of efficiency, although theory predicts this to be the other way round for the situation that we consider.

Stay-or-stay ($w_e = 50$)

Deviations from the theoretical predictions are the largest for the stay-or-stay rule. The actual investment rate is 43% whereas theory predicts a zero investment rate. In line with theoretical predictions, low-productivity workers are almost always assigned the easy job. However, 69% of the high-productivity workers are promoted, where it is predicted that such workers would also be kept in the easy job. In fact, assignment again depends on whether the worker made the investment or not. High-productivity workers that made the investment are promoted in 74% of the cases, those high productivity workers that did not invest (but simply turned out to be lucky) are promoted in only 55% of the cases. This points to the presence of a positive reciprocity mechanism. The employer considers investment a friendly action that justifies a reward in the form of promotion. This reciprocal reaction costs the employer 15 points and yields the worker a benefit of 60 points.

The presence of the positive reciprocity mechanism boosts both investment incentives and efficiency.[6] Given the actual assignment patterns, the expected pay-offs of investment equal 58.3 points and the expected pay-offs from not investing equal 58.25 points. These expected pay-offs are almost identical, explaining why on average workers are indifferent between investment and no investment. With respect to efficiency, under stay-or-stay there are actually efficiency gains on two accounts. One is due to the higher than predicted investment rate, the other to the higher than predicted promotion rate of high-productivity workers. On the other side of the balance sheet is a small loss due to the 3% low-productivity workers not assigned to the easy job. Taken together, the realized surplus is on average almost 126 while only 118¾ is predicted.

Promotion rules and the scope for reciprocity

From the above we conclude that the promotion rules differ in the extent to which they give scope to reciprocity. Theory predicts that

under up-or-out the worker will invest. If he does so, there is no reason for the employer to give a reward, because investment is in the worker's self-interest. Consequently, there is no incentive to improve upon the predicted up-or-out assignment of workers, and hence there is no mechanism supporting low-productivity workers being kept within the firm.

In contrast, the up-or-stay and stay-or-stay promotion rules allow for reciprocity in a way that is efficiency enhancing. The theoretical prediction is that the worker will not invest. The two rules provide different incentives to reciprocate. Under up-or-stay the worker may decide to invest because he anticipates that the employer will punish non-investment by offering a less attractive job than corresponds with his productivity. This anticipated negative reciprocity mechanism does indeed operate. In case of stay-or-stay positive reciprocity plays a role. The worker may want to invest if he anticipates that the firm will reward that by offering a more attractive job. Also this mechanism appears to be present, thereby increasing investment levels and also improving assignment efficiency.

For employer reciprocity to have an impact on investment incentives, the worker must be able to anticipate this kind of behaviour. In the experiment workers could do so, because they had ample opportunities to experience employers' average behaviour during the 30 rounds that were played. However, because interaction was anonymous and the matching changed over the rounds, workers were not informed about the track records of individual employers. Clearly, in practice individual learning opportunities are different. On the one hand learning opportunities are more limited, because workers typically experience only a few different employers themselves. On the other hand, in reality the track record of individual employers can be identified; potential employees can learn from the experiences of former and current employees of a particular employer. Because the firm is a 'long-run' player with potentially many workers, it then has the opportunity (and the incentive) to build up a reputation for being reciprocal.

8.3 Overview of other experimental studies

The finding that in practice underinvestment appears less of a problem than theory predicts it to be is robust; it has been observed by various scholars and in a variety of experimental settings where underinvestment

is predicted to play an important role.[7] These studies confirm that a partial solution to underinvestment is provided by reciprocity (and fairness) considerations. Because investment is typically considered to be fair behaviour, agents reward investors by granting them a higher than predicted return. Given the existence of this informal reciprocity mechanism, there is less scope for efficiency improvement than theory predicts.[8] The need for contractual solutions to underinvestment thus may be substantially less strong in practice.

Despite these findings, it is of interest to verify whether (contractual) solutions to underinvestment work as predicted and do have the (dis)advantages attributed to them by theory. In various related experiments we have considered a range of proposed solutions to underinvestment. Apart from the promotion policies discussed in section 8.2, these include elaborate contractual solutions based on restructuring of the ex post bargaining process, breach remedies, and disclosure rules and privacy rights. In the experiments the focus is typically on the *underlying mechanism* of the proposed solution. The idea is that when the underlying mechanism is not supported by the data, it is very unlikely that the solution based on it will work in practice. Moreover, this approach allows the use of tests based on comparative statics rather than on point predictions. This yields conclusions which are much more robust. We now briefly discuss some of the results obtained.

A large class of contractual solutions to underinvestment relies on the relationship between specific investments and the so-called *outside option principle*. The latter predicts that when two parties bargain over the division of a surplus, the party with an attractive (i.e. binding) outside option will get a share that exactly matches this outside option.[9] The other party then becomes residual claimant, creating efficient investment incentives. In Sloof et al. (2004) we find no support for the predicted comparative statics relationship. When outside options are binding underinvestment occurs, while it is less of a problem than predicted when the outside option is non-binding. The latter can partly be explained by a self-serving bias. Because non-investors have the outside option available, investors feel entitled to a larger return on investment than they can actually get out of the bargaining (and also larger than theoretically predicted). Overall these results suggest that contractual solutions that rely on the outside option principle are unlikely to solve underinvestment.

In practice labour contracts sometimes incorporate a remedy for breach of contract. In case the worker leaves, he has to pay a certain amount to reimburse the employer for the training expenses paid. In that way payback clauses can protect relationship-specific investments. In general there are various ways of calculating the amount of damages that has to be paid. In Sloof et al. (2003, 2006) we evaluate the breach remedies that are most commonly used in practice. These are: *liquidated damages* (the initial contract specifies the exact amount the worker has to pay), *expectation damages* (the worker compensates the employer such that the latter is equally well off as under contract performance) and *reliance damages* (the worker makes the employer equally well off as before the contract had been signed; in practice this implies that the worker pays back the direct costs of training).

Standard theory predicts that the above remedies are typically overzealous in protecting specific investments, as they are expected to induce substantial overinvestment. Our experiments reveal that the predicted motives to overinvest are indeed present, although negative reciprocity acts as a minor counteractive force. We generally observe slightly less overinvestment than predicted. Breach remedies thus do appear to have the disadvantages ascribed to them in the theoretical literature. Moreover, there is much less need for sophisticated breach remedies than theory suggests. Parties are typically better off either by having no payback clause at all (because reciprocity mitigates underinvestment), or by adopting a contract that simply forbids unilateral separation (so-called specific performance contracts).

Finally, theory suggests that the underinvestment problem may disappear when the investor has private information about, for example, the actual return to investment or about the value of outside options. The underlying idea is that private information creates an informational rent in the ex post bargaining, boosting the ex ante investment incentives (when this rent is increasing in the investment made). In Sloof et al. (2002) we look at a situation in which the investor is privately informed about the actual returns to investment.[10] In line with theoretical predictions private information boosts investment incentives when there is limited scope for fairness and reciprocity. But with sufficient scope for these motivational factors, i.e. when the potential returns to investment are high relative to the costs of investment, unobservability does not affect specific investments. Hence only under a restricted set of circumstances can privacy rights be used as an

effective instrument against underinvestment. Sloof (2003) considers the situation where parties are privately informed about their outside options. It is observed that the informal reciprocity mechanism that alleviates underinvestment also carries over to this more realistic case.

The general picture that emerges from all these findings is that some of the theoretically suggested solutions to underinvestment appear to work (e.g. breach remedies), while others do not (e.g. solutions that rely on the outside option principle). Moreover, the disadvantages theoretically attributed to some solutions to underinvestment (e.g. assignment inefficiency in the case of promotions, overinvestment in the case of breach remedies) are typically observed in the laboratory.

8.4 Conclusion and discussion

The typical objection raised against experimental results is that they are obtained in an artificial laboratory setting and that in reality people will behave differently. One criticism that is often made is that the stakes in laboratory experiments are usually rather low compared to those in the real world. Experiments that explicitly address this issue by increasing the stakes to economically highly significant amounts find only a weak effect on behaviour; higher stakes do not affect the main conclusions. In particular, motivational factors like fairness and reciprocity continue to play an important role; see Camerer (2003) for an overview.

Another usual criticism is that the experiments are typically conducted with students. This is problematic when the behaviour of student subjects is not representative of the behaviour of those taking investment decisions in the real world (i.e. actual workers and employers). Fehr and List (2004) explicitly address this issue in an experimental setting considering underinvestment. They do find some differences between students and a subject pool of chief executive officers, but not in a direction one a priori would expect; CEOs invest *more* and behave *more* reciprocally than student subjects. If anything, one would thus conclude that fairness and reciprocity motivations are even more important in practice.

A third potential criticism is that in the experiments reviewed in this chapter subjects were anonymously paired and could not communicate with each other. In reality interactions are typically not anonymous and parties are able to communicate. Ellingsen and Johannesson (2004b) explicitly investigate the impact of communication in a setting

with potential underinvestment. In their experiment subjects could make non-binding promises and threats. Theoretically these cheap talk messages should not have an impact, yet it appears that these promises and threats do mitigate the underinvestment problem. It can similarly be expected that under repeated, non-anonymous interaction, as typically applies in practice, people are less likely to behave opportunistically. Economic agents might then build a reputation for trustworthy behaviour.

Notwithstanding the above remarks, it remains important to verify the external validity of the experimental results: do the findings obtained in the laboratory carry over to the real world? In the absence of information about this, conclusions from experiments remain premature. Compared to the existing studies future research thus should focus more on assessing external validity. Such an assessment entails that the insights obtained from experiments are incorporated into, and verified by means of field studies. This is far from easy, given the serious problems with obtaining field data that induced the use of the experimental method in the first place. Sometimes careful econometric techniques can provide a way out though. For example, as explained in the introduction, our first main conclusion that underinvestment is less of a problem than theory predicts is well in line with the small selectivity-corrected estimates of the wage returns to private sector training reported in chapter 3. Taking the results from both field and experimental data together, then, the empirical case for underinvestment in firm-specific human capital is rather weak. Recently some interesting results have also been obtained on the importance of reciprocity in practice, which we will briefly discuss (see also section 3.6).

In Leuven et al. (2005) we use a survey held in 2001 among a representative sample of the Dutch population aged 16–64. Besides background characteristics like formal education, age and gender, this dataset also contains information about participation in work-related training and who (employer or worker) paid for this training. By incorporating an additional survey question, we try to measure the reciprocal attitude of employees. In particular, we ask them: 'If someone does something that is beneficial to you, would you be prepared to return a favour, even when this was not agreed upon in advance?' The answers given enable a classification of employees into three different categories, i.e. employees having either weak, intermediate or strong reciprocal motivations. If reciprocity really matters one would expect,

based on the existing experimental findings, that the more reciprocal workers are more likely to obtain firm-sponsored training. And this is exactly what we observe. Workers with a high sensitivity to reciprocity have a 15 percentage points higher firm-sponsored training rate than those with a low sensitivity to reciprocity. This result is robust to differences in background characteristics between workers. Interestingly, participation in training that is fully paid for by the worker her/himself is unrelated to her/his reciprocity type, as one would expect. All in all the key role of reciprocity for underinvestment thus appears to carry over to the real world and employers seem to actively rely on reciprocity as an informal enforcement mechanism.

The results reported in Leuven et al. (2005) are promising, because they indicate that some important insights obtained from experiments are indeed highly relevant in practice. More work should be done though, before we can convincingly translate the findings into valuable policy implications. Yet the results up till now do suggest that there does not seem much need for government intervention, nor for firms to change their HRM policy. Underinvestment is much less of a problem than it is predicted to be. And, besides that, employers seem to actively rely on (informal) mechanisms against it.

Notes

1. Small wage returns indicate that the marginal returns to additional investments are small, and thus provide no evidence for substantial efficiency losses due to underinvestment.
2. In the experiment subjects earn points on the basis of the *actual* choices they make. At the end of the experiment these points are converted into euros. The experimental points reflect the economic costs and benefits of the various decisions. The numbers mentioned in the main text are the ones used in the experiment.
3. The worker's outside productivity is taken as the benchmark and normalized to zero. Productivities within the firm are specified in comparison to productivity elsewhere. The important point to note is that the worker's productivity within the firm is always higher than elsewhere.
4. The difference between 220 (i.e. productivity of a high type of worker in the difficult job) and 175 (productivity of a high type of worker in the easy job) can be interpreted as the extent to which the two jobs really differ. When these numbers become very close, the firm is simply generating job titles. Although in the experiment we keep the various productivity levels

fixed (as in table 8.1), the stylized setting with only two different jobs can thus capture various types of job designs within firms.

5. In fact it can be shown that no wage combination (w_d, w_e) does so in the setting that we consider. Theory thus predicts that the trade-off between inefficient investment and inefficient matching is unavoidable.
6. Positive reciprocity obviously only applies to workers who did make the investment. The 55% promotion rate amongst the high-productivity workers who did not invest can be explained by anticipated negative reciprocity. A high-productivity worker who is not promoted may feel mistreated. He may then punish the employer by not accepting the easy job offered. As reported in Oosterbeek et al. (2006), 46% of the high-productivity workers who are offered the easy job do not accept this offer and quit. Employers that anticipate this reciprocal response are indeed better off by granting the worker promotion. In fact, an additional treatment in which the worker has private information about the investment made reveals that anticipated negative reciprocity is (in this setting) a stronger mechanism than positive reciprocity.
7. See, for example, Berg et al. (1995), Ellingsen and Johannesson (2004a, 2004b), Gantner et al. (2001), Hackett (1993, 1994), Königstein (2000), Oosterbeek et al. (2001, 2003), Sloof et al. (2002, 2006) and Sonnemans et al. (2001).
8. Other recent experimental studies reveal that reciprocity also has an impact on contractual choices (Bohnet et al. 2001; Fehr et al. 2004; Fehr and Schmidt 2000). The main result of these papers is that many contracts are deliberately left incomplete in order to rely on reciprocity as the more powerful enforcement device (see also Fehr et al., 1997).
9. Here the term 'outside option' should not be taken literally; it in fact refers to any option that, if taken, effectively puts an end to the bargaining. In the proposed solutions such 'outside options' are *endogenously* created through the design of the initial contract. One could, for example, think of a one-sided option to extend the contract for some additional period of time; this kind of provision is not uncommon in contracts of soccer players.
10. In the experiment described in section 8.2 this would correspond to a situation in which only the worker knows his own productivity.

References

Berg, J., J. Dickhaut and K. McCabe (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10: 122–42.

Bewley, T. F. (1999). *Why Wages Don't Fall during a Recession*. Cambridge, MA: Harvard University Press.

Bohnet, I., B. S. Frey and S. Huck (2001). More order with less law: on contract enforcement, trust and crowding. *American Political Science Review*, 95: 131–44.

Camerer, C. F. (2003). *Behavioral Game Theory. Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.

Ellingsen, T. and M. Johannesson (2004a). Is there a hold-up problem? *Scandinavian Journal of Economics*, 106: 475–94.

(2004b). Promises, threats and fairness. *Economic Journal*, 114: 397–420.

Fehr, E. and J. A. List (2004). The hidden costs and returns of incentives – trust and trustworthiness among CEOs. *Journal of the European Economic Association*, 2: 743–71.

Fehr, E. and K. Schmidt (2000). Fairness, incentives, and contractual choices. *European Economic Review*, 44: 1057–68.

Fehr, E., S. Gächter and G. Kirchsteiger (1997). Reciprocity as a contract enforcement device: experimental evidence. *Econometrica*, 65: 833–60.

Fehr, E., A. Klein and K. Schmidt (2004). Fairness, contracts and incentives. CEPR Discussion Paper 4464.

Gantner, A., W. Güth and M. Königstein (2001). Equitable choices in bargaining games with advance production. *Journal of Economic Behavior and Organisation*, 46: 209–25.

Hackett, S. (1993). Incomplete contracting. A laboratory experimental analysis. *Economic Inquiry*, 31: 274–97.

(1994). Is relational exchange possible in the absence of reputations and repeated contract? *Journal of Law, Economics and Organization*, 10: 360–89.

Königstein, M. (2000). *Equity, Efficiency and Evolutionary Stability in Bargaining Games with Joint Production*. Berlin: Springer Verlag.

Leuven, E., H. Oosterbeek, R. Sloof and C. van Klaveren (2005). Worker reciprocity and employer investment in training. *Economica*, 72: 137–49.

Malcomson, J. M. (1997). Contracts, holdup, and labor markets. *Journal of Economic Literature*, 35: 1916–57.

(1999). Individual employment contracts, in, O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, vol. IIIB. Amsterdam: North-Holland, pp. 2292–2372.

Milgrom, P. and J. Roberts (1992). *Economics, Organization and Management*. Prentice-Hall: Englewood Cliffs, NJ.

Oosterbeek, H., R. Sloof and J. Sonnemans (2001). Who should invest in specific training? forthcoming in *Journal of Population Economics*.

(2006). Promotion rules and skill acquisition: an experimental study. *Economica*, forthcoming.

Oosterbeek, H., J. Sonnemans and S. van Velzen (2003). The need for marriage contracts: an experimental study. *Journal of Population Economics*, 16: 431–53.

Sloof, R. (2003). Price setting power versus private information: an experimental evaluation of their impact on holdup. SCHOLAR Working Paper Series WP 42/04, Amsterdam.

Sloof, R., H. Oosterbeek and J. Sonnemans (2002). Does making investments unobservable boost investment incentives? Amsterdam. Forthcoming in *Journal of Economics and Management Strategy*.

Sloof, R., E. Leuven, H. Oosterbeek and J. Sonnemans (2003). An experimental comparison of reliance levels under alternative breach remedies. *RAND Journal of Economics*, 34: 205–22.

Sloof, R., J. Sonnemans and H. Oosterbeek (2004). Specific investments, holdup, and the outside option principle: an experimental study. *European Economic Review*, 48: 1399–1410.

Sloof, R., H. Oosterbeek, A. Riedl and J. Sonnemans (2006). Breach remedies, reliance and renegotiation. *International Review of Law and Economics*, forthcoming.

Sonnemans, J., H. Oosterbeek and R. Sloof (2001). On the relation between asset ownership and specific investments. *Economic Journal*, 111: 791–820.