

Altmann, Steffen; Falk, Armin; Huffman, David

Working Paper

Implicit contracts, unemployment, and labor market segmentation

IZA Discussion Papers, No. 5001

Provided in Cooperation with:

IZA – Institute of Labor Economics

Suggested Citation: Altmann, Steffen; Falk, Armin; Huffman, David (2010) : Implicit contracts, unemployment, and labor market segmentation, IZA Discussion Papers, No. 5001, Institute for the Study of Labor (IZA), Bonn

This Version is available at:

<https://hdl.handle.net/10419/36857>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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

IZA DP No. 5001

Implicit Contracts, Unemployment, and Labor Market Segmentation

Steffen Altmann
Armin Falk
David Huffman

June 2010

Implicit Contracts, Unemployment, and Labor Market Segmentation

Steffen Altmann

IZA and University of Bonn

Armin Falk

University of Bonn and IZA

David Huffman

Swarthmore College and IZA

Discussion Paper No. 5001
June 2010

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Implicit Contracts, Unemployment, and Labor Market Segmentation*

We analyze the impact of imperfect contract enforcement on the emergence of unemployment. In an experimental labor market where trading parties can form long-term employment relationships, we compare a work environment where effort is observable, but not verifiable to a situation where explicit contracts are feasible. Our main result shows that unemployment is much higher when third-party contract enforcement is absent. Unemployment is involuntary, being caused by firms' employment and contracting policy. Moreover, we show that implicit contracting can lead to a segmentation of the labor market. Firms in both segments earn similar profits, but workers in the secondary sector face much less favorable conditions than their counterparts in primary-sector jobs.

JEL Classification: C92, J64, M55

Keywords: incentives, implicit contracts, unemployment, fairness, dual labor markets

Corresponding author:

Steffen Altmann
IZA
P.O. Box 7240
53072 Bonn
Germany
E-mail: altmann@iza.org

* Financial support from the Deutsche Forschungsgesellschaft through SFB/TR 15 is gratefully acknowledged. We thank conference and seminar participants in Alicante, Bonn, Buch, Graz, Maastricht, Milan, Pittsburgh, Rotterdam, and Zurich for valuable comments.

1 Introduction

Many employment relationships are characterized by contractual incompleteness. Probably the most important reason for leaving contracts incomplete is the difficulty to monitor and verify employee performance or work effort. When this is the case, firms cannot stipulate fully contingent contracts but have to rely on other, implicit means to elicit work effort. Several such implicit contract enforcement mechanisms have been discussed in the literature, ranging from voluntary bonus payments to self-enforcing relational contracts based on an implicit threat to dismiss workers who have been caught shirking (e.g., MacLeod and Malcomson 1989, Baker, Gibbons, and Murphy 1992, MacLeod and Malcomson 1998). Other papers have argued that social preferences can generate the implicit incentives necessary to overcome moral hazard problems. Workers who feel treated fairly, e.g., because their employer generously shares the rents from production, might voluntarily provide high work effort in return (Akerlof 1982, Fehr, Gächter, and Kirchsteiger 1997).

While implicit contracts can provide strong performance incentives, they are potentially detrimental to the labor market as a whole. In particular, implicit contracting can give rise to involuntary unemployment. Firms might find it optimal to dismiss workers rather than to cut wages in times of economic downturns if they fear adverse reactions to pay cuts (Bewley 1999). The requirement to pay strictly positive rents to workers can also render less productive jobs unprofitable and therefore lead to job rationing. Finally, unemployment could itself be a prerequisite to encourage effort provision of employed workers. Higher unemployment increases the cost of job losses since, c.p., the job acquisition rate for unemployed workers decreases. Thus, unemployment has been argued to provide performance incentives for employed workers who want to avoid losing their jobs (Shapiro and Stiglitz 1984, MacLeod and Malcomson 1989).

In this paper, we empirically analyze the relationship between the degree to which contracts are explicitly enforceable and the emergence of unemployment. In order to identify the causal impact of contract enforcement on the level of unemployment, we implement an experimental labor market where we exogenously vary the verifiability of work effort. In the market, firms and workers interact during multiple market periods. All firms share the same

production technology which exhibits decreasing returns to scale from labor, but ensures that full employment is technologically efficient. In our main treatment (*IC treatment*), firms can observe but not verify their workers' effort. Firms thus have to rely on implicit incentives to elicit work effort. In a control treatment, concluded contracts are explicitly enforced; i.e., a worker's effort must be equal to the contractually agreed upon effort level (*C treatment*). In other words, effort is verifiable and there is no contract enforcement problem in this treatment.

Our main result shows that the absence of third party contract enforcement has a dramatic impact on the level of unemployment. While there is virtually zero unemployment when contracts are explicitly enforced, unemployment is much higher when the contracting parties have to rely on implicit performance incentives. We show that the difference in unemployment is attributable to differences in how labor markets function under explicit and implicit contract enforcement. When third party contract enforcement is not feasible, firms rely on a mixture of instruments to generate implicit performance incentives. They pay strictly positive rents to their workers, and many firms try to build up long-term employment relations. In such relational contracts, high-performing workers are rewarded by being re-employed and earning positive rents in future periods; shirking workers are punished through dismissal. Moreover, many firms ration jobs and offer fewer vacancies than feasible. Job rationing together with the presence of high worker rents indicates that unemployment in the IC treatment is involuntary. In the C treatment where effort is verifiable and explicitly enforced, labor market outcomes differ substantially along all these dimensions. Firms pay wages close to the market clearing level and reap the major share of production surplus. Employment relations are shorter than in the IC treatment and the overwhelming majority of firms does not ration jobs. As a result, unemployment in this treatment is very low and mostly voluntary, being caused by workers who do not accept existing contract offers.

The necessity to rely on implicit performance incentives also generates strong market dynamics in the IC treatment. Since workers condition their work effort on the wage received, firms generally earn higher profits when paying generous rents to workers. At the same time, the decreasing-returns-to-scale production technology generates a potential trade-off between rents paid per worker, the number of workers employed, and firm profits. Indeed, in early

periods of the experiment we find that firm profits are highest when generous rent payments are accompanied with a rationing of job offers through firms. The difference in profits between firms who do and those who do not ration jobs creates strong dynamics towards adopting a strategy that involves job rationing; this in turn yields a substantial increase in the level of unemployment. After an initial phase in which we observe this trend, however, a plateau seems to be reached and a relatively stable fraction of firms continues to operate without rationing job offers. In this later phase, unemployment stabilizes on a high level and firm profits do no longer depend on firm size. Rather, two types of firms seem to coexist in the long run. We show that employment relationships in the different firm types differ in several important dimensions other than firm size. In particular, workers earn substantially lower rents and we observe higher worker turnover in firms which do not ration jobs. In the long run, the situation in the IC treatment thus resembles a segmented labor market in which some workers are employed in “primary sector” jobs involving relatively stable employment relationships and generous rent sharing and other workers are working under less favorable conditions in “secondary sector” jobs (Doeringer and Piore 1971, Saint-Paul 1996).

Efficiency wage theories have long postulated that the absence of explicit contract enforcement can lead to involuntary unemployment (Shapiro and Stiglitz 1984, MacLeod and Malcomson 1989, Akerlof and Yellen 1990). In this paper we empirically establish a direct causal link between contract enforcement and the emergence of unemployment. So far, there has been only indirect evidence which suggested that this link might exist. For instance, in labor market surveys as well as in economic experiments it has been shown that fairness concerns in environments characterized by contractual incompleteness can be the cause of wage rigidities (Campbell III and Kamlani 1997, Bewley 1999, Fehr and Falk 1999). In a setup similar to ours, Brown, Falk, and Fehr (2004) have shown that the absence of third party contract enforcement leads to a “bilateralization” of trade, where interaction between firms and workers is characterized by long-term employment relationships with high wages and high effort. Unemployment, however, is exogenously given in their experiments and firms have no option to ration jobs.

The paper most closely related to ours is the work by Fehr, Kirchsteiger, and Riedl (1996) who study imperfect monitoring and unemployment in experimental labor markets. How-

ever, their experimental design differs from ours in a variety of important aspects. First, they concentrate on one-shot interactions between firms and workers while our focus is on relational contracts. Our results show that this feature is important in the absence of explicit contract enforcement since it allows firms and workers to make use of the implicit incentives generated through repeated interaction. Second, as a shortcut for implementing a threat of dismissal, Fehr, Kirchsteiger, and Riedl (1996) introduce an exogenously given, strictly positive probability that workers who shirk will be caught and have to pay a penalty to their firm. Our setup allows us to directly observe dismissal and re-employment decisions of firms. Finally, workers' productivity in their experiment is private knowledge to firms and productivity of some firms is so low that these firms cannot profitably offer incentive compatible contracts. In contrast, we find that firms ration jobs and involuntary unemployment emerges *although* it is always possible for firms to offer an incentive compatible contract that yields positive profits.

Our finding of an endogenous segmentation of the labor market in the IC treatment also provides interesting new insights for understanding dual labor markets. It has long been acknowledged that dual labor markets can be an implication of contract enforcement problems and efficiency wages. However, in previous models of market segmentation, dual labor markets typically arise as a consequence of differences in monitoring technology across sectors (Bulow and Summers 1986) or due to differences in adjustment costs to demand fluctuations (Saint-Paul 1996, ch. 4). In our experiment segmentation occurs in the IC treatment even though all firms face the same technological constraints. Rather, segmentation is attributable to complex market feedback mechanisms under contractual incompleteness. In response to the non-verifiability of work effort many firms ration jobs, pay high wages and employ specific workers over a long time horizon. As a result, unemployment increases which in turn allows other firms to profitably employ a “secondary-sector” contracting strategy, involving lower wage payments and tighter conditions for contract renewal. We show that—in the presence of high unemployment—workers in these firms become increasingly willing to provide effort for a given wage. This holds both with respect to their own behavior earlier on in the experiment and with respect to workers who are mainly employed in primary sector firms. It is this change in behavior of secondary-sector workers that drives the convergence

in firm profits and, ultimately, enables the coexistence of two different firm strategies in the market.

The remainder of the paper is organized as follows: the following section outlines our experimental setup. Section 3 derives behavioral predictions and Section 4 presents the empirical results. Section 5 concludes.

2 Experimental Design and Procedures

To study the impact of contractual incompleteness on unemployment, we implemented an experimental labor market where we exogenously varied the verifiability of work effort. In the market, firms and workers interacted during 18 market periods. Each of the 18 periods consisted of two stages: a market phase where firms offered employment contracts and hired workers, and a work phase where work effort of employed workers was determined. The experimental treatments differed only in the degree to which work effort in the second phase was third party enforceable. In our main treatment, henceforth called *Incomplete Contracts Treatment* (or *IC treatment*), third party contract enforcement was absent and workers thus could depart from the contractually agreed upon effort level. By contrast, the effort level stipulated in the employment contract was explicitly enforced in our control treatment, henceforth called *Complete Contracts Treatment* (or *C treatment*). Keeping everything else identical (production technology, supply and demand of labor, etc) while varying the level of third party enforcement allows us to causally identify the effects of contractual incompleteness. This provides complementary evidence to non-experimental studies of labor market outcomes (e.g., Krueger and Summers 1988, Blanchflower, Oswald, and Sanfey 1996; see Falk and Heckman 2009 for a methodological discussion).

2.1 The Market Phase

Firms were the contract makers in the market phase. When offering a contract, firms stipulated a non-contingent wage payment w and a desired level of effort \hat{e} . To study the relevance of long-term employment relationships in the different treatment conditions, firms could make two types of contract offers: public offers which were available to all workers and could also be observed by all other firms, or private contract offers that were only available to

one specific worker. The latter type of contract offer allowed firms to rehire certain workers and interact repeatedly with them. To enable the formation of long-term relations, in the beginning of the experiment each worker and each firm received an identification number (ID) which was held constant throughout the whole experimental session. If an employer wanted to (re)hire a specific worker via a private contract offer, she had to specify the ID of the worker in addition to the wage and desired effort level when entering the contract offer. In this case, only the selected worker was informed about the contract offer, and only this worker could accept the offer.

In a certain market period, each employer could hire up to two workers. As long as none of her contract offers had been accepted in a given period, an employer could make as many private and public offers as she wanted. A worker could accept all contract offers available to him, i.e., all public offers that were not yet accepted and private offers that firms had addressed to him. Once a worker accepted a contract offer, the contract between this worker and the respective firm was concluded. After concluding a contract, a worker was not allowed to accept further contract offers in this period. Additionally, all other outstanding offers of the respective employer were removed from the list of available contracts in the moment where one of her contract offers was accepted. The employer could then decide to hire a second worker by entering new contract offers. This market feature was implemented to prevent that an employer who wanted to employ only one worker but entered multiple contract offers had two offers accepted before being able to withdraw her remaining contract offers.

The market phase ended when the maximum number of contracts had been concluded or when all firms had indicated that they did not want to make additional contract offers.¹ At the end of the market period, the worker(s) of a given firm were informed about the contracts concluded by their firm, i.e., each worker received a summary of his own contract terms as well as information on whether and under which conditions his firm had employed

¹We also implemented a maximum trading time of 200 seconds for each market phase. This constraint was, however, only binding in few occasions (mostly in the C treatment). The impact of the time constraint on the level of unemployment and other market outcomes reported below is therefore limited and confined to the control treatment with explicit contract enforcement (see Section 4.1).

a second worker. Providing this information can be important for workers when contracts are not third party enforceable (IC treatment). If some workers base their work effort on the extent to which their firm shares production rents, e.g., because they respond to fair wage payments, knowledge of offered rents, productivity, and firm size is important for workers (see Section 2.3 and Section 3).

2.2 The Work Phase

After the end of the market phase, employed workers entered the second stage of a market period—the work phase. In this stage, actual work effort e was determined. Since effort was contractible in the complete contracts treatment, workers who had accepted a contract offer in this treatment had to comply with the contract terms. The desired effort level \hat{e} stipulated in their contract was thus explicitly enforced; i.e., $e = \hat{e}$ was exogenously implemented by the experimenter. By contrast, work effort was observable, but not verifiable in our main treatment (IC treatment). Therefore, a worker could choose any feasible level of effort in the work phase, i.e., he could also exert less or more effort than stipulated in his employment contract. Workers’ effort choices, together with firms’ wage payments, determined material payoffs of firms and workers. Before the next period started, a firm and its worker(s) were informed about actual work efforts and the resulting payoffs for the firm and the workers employed by this firm.

2.3 Parameters and Procedures

Participants’ roles were randomly assigned at the beginning of the experiment and kept constant throughout all market periods. In every market, we had 17 workers and 7 firms. Since firms could employ at most two workers, this implies that three workers were “exogenously” unemployed in each period (see section 3). A worker’s material payoff π_W was given by

$$\pi_W = \begin{cases} w - c(e) & \text{if worker accepted a contract } [w, \hat{e}] \\ 0 & \text{if unemployed} \end{cases}$$

A worker who remained unemployed in a given period received a payoff of 0 points. An employed worker received the wage w specified in his contract and had to bear the cost of the work effort he provided, $c(e)$. The set of feasible efforts and wages was given by

$e \in \{1, 2, \dots, 10\}$ and $w \in \{0, 1, 2, \dots, 100\}$. Effort costs $c(e)$ increased convexly in the level of actual work effort (see Table 1).

Effort level e	1	2	3	4	5	6	7	8	9	10
Cost of effort $c(e)$	0	1	2	4	6	8	10	12	15	18

Table 1: *Schedule of effort costs.*

A firm's material payoff depended on the number of workers hired, the wage(s) paid, and the effort exerted by the worker(s). Firms' production technology was characterized by decreasing returns to scale. Specifically, each unit of effort by a worker increased production (and the firm's payoff) by 10 points if only one worker was employed by the firm. If two workers were employed, each unit of effort increased the firm's payoff by 7 points. The contractually stipulated wage payments diminished firms' profits. The material payoff of a firm, π_F , can therefore be summarized as follows:

$$\pi_F = \begin{cases} 10e_1 - w_1 & \text{if one worker employed} \\ 7(e_1 + e_2) - w_1 - w_2 & \text{if two workers employed} \\ 0 & \text{else} \end{cases}$$

e_1 (e_2) denotes the effort provided by the first (second) worker, and w_1 (w_2) is the wage paid to the first (second) worker employed by the firm. Note that this specification of the production technology implies that efficiency is maximized when two workers are employed and maximum effort is exerted: the second worker's marginal productivity per unit of effort is 4 whereas the marginal cost of effort lies between 1 and 3 points. Payoff functions π_F and π_W , workers' cost schedule $c(e)$ as well as the number of firms and workers in the market were common knowledge.

The experiment was carried out in the BonnEconLab, the laboratory for economic experiments at the University of Bonn. We conducted five sessions each for the IC treatment and the C treatment. A total of 240 subjects, mainly undergraduate university students from all majors, took part in the experiments. Every subject participated only in one of the treatment conditions. At the beginning of an experimental session, participants received detailed

information about the rules and structure of the experiment.² The experiment started only after all participants had answered several control questions correctly. In addition, subjects played one trial period of the market phase to ensure that they understood how to use the computer program. Sessions lasted about 110 minutes and subjects earned on average 25.49 Euro (about 35 USD at the time of the experiment), including a showup fee of 8 Euro. The experiments were computerized using the software “z-Tree” (Fischbacher 2007); subjects were recruited with the online recruitment system by Greiner (2003).

3 Behavioral Predictions

As indicated in Section 2, efficiency is maximized when all firms employ two workers and all workers exert maximum effort. If it is common knowledge that players are rational and selfish, and if effort is contractible (C treatment), we should expect that profit-maximizing firms implement the first best outcome and reap all gains from trade: in each period, firms will employ two workers, offering contracts which ask for maximum effort ($\hat{e} = e = 10$) and pay the minimum acceptable wage. As unemployed workers receive zero unemployment benefits, this wage is equal to the cost of the implemented effort, i.e., $w = c(10) = 18$ points. Since there are 7 firms and 17 workers in each market, three workers will remain unemployed by design of the experiment. As we are interested in unemployment which arises *endogenously* through market interaction between firms and workers, we will refer to the situation where the maximum number of workers is employed and only three workers remain “exogenously” unemployed as *full employment* in what follows.

Maintaining the benchmark assumption of materially selfish agents, the absence of third party contract enforcement is predicted to lead to different outcomes in terms of wage offers and effort choices, but not unemployment. Since our setup constitutes a game of finite length and since effort is costly, but not verifiable, workers will choose the minimum effort $e = 1$ in the final period of the IC treatment, irrespective of the wage and desired effort level stipulated in their contract. Anticipating this, money-maximizing firms will pay the

²A translation of the instructions is available upon request. To rule out that differences in participants’ experiences from their employment relations outside the laboratory could bias our results, instructions were framed in a neutral goods-market language.

minimum acceptable wage for an effort of 1 in this period. This amounts to a wage payment $w = c(1) = 0$ since workers' outside option is to remain unemployed and receive zero rents.³ Through backward induction, the outcome of minimal effort (and minimal wage) should also hold in all pre-final periods of the IC treatment. This contrasts with non-minimal wages and high effort levels when contracts are complete. Importantly, however, the two treatments should *not* differ in the second dimension of market efficiency—the level of unemployment. In particular, under the assumption of money-maximizing behavior of all players we should still expect full employment even though contracts are not third party enforceable. To see this, note that—given our assumptions on production technology and payoff functions—it is more profitable for a firm to employ two workers at the minimum possible wage ($w = 0$) who exert minimum effort ($e = 1$) compared to employing just one worker at such a wage-effort combination.

If the assumption of all selfish types does not hold, and some agents are fair-minded in the sense of losing utility when rents are shared unequally, the predictions for the complete contract setting are largely unchanged. Fairness might lead some workers to reject offers that involve very unequal payoffs in the C treatment, but theory and previous evidence suggest that this is unlikely to happen (see, e.g., Fehr and Schmidt 1999). In particular, the competition for jobs in our setting means that workers cannot effectively punish a firm by rejecting an offer; a firm can likely just find another worker. Without the potential for equalizing payoffs, rejection does not improve utility of fair workers, and in such circumstances even fair-minded workers tend to accept contracts with unequal rents. Thus, firms are still predicted to achieve close to the first best, and workers are predicted to earn zero rents.

With incomplete contracts, by contrast, the presence of fair types substantially changes predictions. Rather than expressing fairness concerns by rejecting offers, fair workers can accept a contract and then equalize payoffs through the effort choice. As a consequence, it could be that effort levels are an increasing function of the wage offer, because fair workers reciprocate a “fair” wage by choosing a high effort level. Indeed, there is mounting evidence both from laboratory (e.g., Fehr and Gächter 2000, Brown, Falk, and Fehr 2004) and field

³Assuming that workers reject contract offers with net payoffs equal to their outside option yields $[w, \hat{e}] = [1, 1]$ in IC and $[w, \hat{e}] = [19, 10]$ in C.

studies (e.g., Bewley 1999, Kube, Maréchal, and Puppe 2008, Cohn, Fehr, and Götte 2008) which suggests that the slope of $e(w)$ is positive due to some workers having fairness concerns (see also Akerlof 1982 and Akerlof and Yellen 1990). Whether it is actually profitable for firms to offer positive rents, however, depends crucially on the magnitude of the effort response. In a repeated interaction setting like ours, it is plausible that strategic behavior of money-maximizing workers actually strengthens the responsiveness of effort to the wage, simply because selfish workers have an incentive to imitate fair workers if this causes firms to re-hire them and pay rents in the future (Shapiro and Stiglitz 1984, MacLeod and Malcomson 1998). In our setting, firms can engage in such long-term relational contracts by using private contract offers and combining rent payments with an implicit threat of firing in case a worker does not provide sufficiently high effort. Future rents provide an incentive for money-maximizing workers to perform, and current rent payments may also be important, as a signal that the employer will not try to renege on wage payments in the future (MacLeod and Malcomson 1998). Notably, the presence of some fair workers makes it credible that firms will pay non-minimal wages (rents) even in the final market period. This potentially allows firms to provide an incentive for selfish workers to perform in the pre-final and earlier periods, despite the finite horizon of the game.⁴ Thus, due to both fairness and strategic motives of workers, it may be profitable for firms to offer positive rents and to engage in long-term relationships when contracts are incomplete.

The necessity to pay strictly positive rents to elicit work effort under contractual incompleteness, however, might lead to unemployment in the IC treatment. Because of decreasing returns to scale, firms which employ two workers might not be able to offer sufficiently generous rents compared to firms which employ only one worker. It could thus be profitable for firms in the IC treatment to ration job offers and employ only one worker at a generous wage. This, however, gives rise to involuntary unemployment. Workers would be willing to accept

⁴The presence of a final period rent avoids the problem of unraveling due to backward induction. More formally, Brown, Falk, and Fehr (2004) have shown for a labor market similar to our IC treatment that the presence of a minority of fair-minded players is sufficient to sustain a pooling equilibrium where firms build up long-term employment relationships involving strictly positive rent payments and a policy of contingent contract renewal and where all workers provide high levels of effort in pre-final periods of the game. In the final period, fair workers provide payoff equalizing effort, while selfish workers shirk.

contract offers for (less than) the prevailing wage, but firms do not offer enough contracts. Obviously, the relative profitability of employing one or two workers strongly depends on the wage payments that are necessary to induce work effort; i.e., on the slope of $e(w)$. For instance, in the presence of fair-minded workers it is crucial what these workers consider to be a fair wage.

It is possible that what constitutes a sufficiently generous wage from a worker’s perspective is itself influenced by market conditions. This could come about through a changed perception of what is a fair wage, as argued by Akerlof and Yellen (1990). Along these lines, Burks, Carpenter, and Verhoogen (2007) have recently provided survey evidence supporting the view that local labor market conditions might affect workers’ fairness perceptions. Changing market conditions can also reduce shirking for incentive reasons. For example, being fired for shirking becomes more costly for a worker in expected terms as unemployment rises, due to the fall in the chance of finding another job. If such mechanisms are relevant, changes in market conditions will feed back into workers’ behavior. As we will see below, the endogenous increase in unemployment and an associated decline in the job acquisition rate indeed contributes to the segmentation of the labor market in the IC treatment, via a change in workers’ behavior.

In the following section, we test the behavioral predictions with our experimental data. We first analyze the causal impact of third party contract enforcement on the level of unemployment. We then study treatment differences in other market characteristics to analyze whether the channels through which contract enforcement affects unemployment are in line with our hypotheses. Finally, we take a closer look at the IC treatment and study how market outcomes and behavior of firms and workers interact.

4 Results

4.1 Unemployment—Aggregate Market Outcomes

We first turn to our main question: does the extent to which explicit contract enforcement is possible causally impact the level of unemployment? Figure 1 depicts the level of “endoge-

nous unemployment”⁵ for the two treatment conditions. Initially, unemployment rates do not differ significantly between the different market environments (Mann-Whitney U-Test for period 1 observations only, $p = 0.217$)⁶. After the first few periods, however, we observe a strong increase in unemployment in the IC treatment, while the level of unemployment stays close to zero when contracts are third party enforceable. As a consequence, the overall level of unemployment differs substantially across treatment conditions. When effort is not verifiable, the average unemployment rate is higher than 30% while it is only about 5% in the C treatment. In all sessions of the IC treatment unemployment is higher than in any of the C sessions. The difference between treatments is statistically significant (Mann-Whitney U-Test, $p < 0.01$).

Result 1: *We observe strong differences in unemployment between treatment conditions. Under explicit contract enforcement (C treatment), unemployment levels are close to the minimal possible level. When effort is not verifiable (IC treatment), unemployment rises strongly before stabilizing at a relatively high level.*

The absence of third party contract enforcement has a strong and positive impact on the level of unemployment. We have hypothesized in Section 3 that this might be the result of differences in how labor markets function under explicit and implicit contract enforcement. In particular, we have argued that firms in the C treatment should operate most profitably if they hire the maximum possible number of workers and minimize wage costs for a given—explicitly enforced—effort level. We therefore should expect low rent payments to workers and high effort levels in this treatment. Furthermore, since effort is contractually enforceable, firms have no reason to engage in relational contracting, i.e., to (re-) employ specific performing workers via private contract offers. We should thus see more public contract offers and shorter employment spells in comparison to the IC treatment.

⁵Remember that, in both treatments, 3 workers were unemployed “by design” in every session and every period due to excess supply of labor. To measure “endogenous unemployment” we therefore calculate the total number of unemployed workers *minus* 3 and divide by the number of possible jobs (given that each of 7 firms could offer 2 vacancies in every period the number of possible jobs in each market is 14).

⁶All non-parametric tests use session averages as independent observations. Reported p-values are always two-sided.

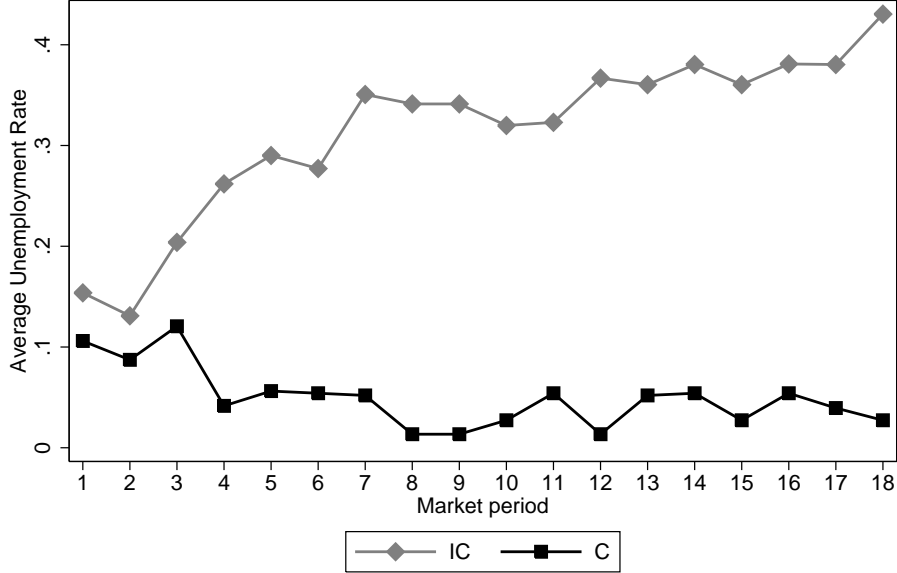


Figure 1: *Evolution of average unemployment over time.*

In contrast to this, firms in the IC treatment have to rely on a combination of generous wages and implicit incentives to elicit work effort. We hypothesized that firms will use private contract offers to implement a policy of contingent contract renewal that combines an (implicit) promise of future rent payments with a threat of contract termination when the worker’s performance is not sufficient. To qualify for future employment, workers are expected to respond to these implicit incentives by reciprocating more generous wage payments with higher performance. We should thus observe a stronger positive correlation between wages and efforts when contracts are not explicitly enforced.

Table 2 shows that our hypotheses on differences in market characteristics are borne out by the data. When contracts are third-party enforceable, markets are characterized by low (offered and realized) worker rents, efforts are close to the maximum level, and firms predominantly hire workers via public contract offers. By contrast, the IC treatment is characterized by more generous wage and rent payments, and efforts and wages are much more strongly correlated.⁷ As hypothesized, long-term employment relationships initiated

⁷Notably, the average wage-effort correlation is about twenty percentage points lower in the final period of IC compared to the overall average correlation, indicating that removal of firing threat affects average behavior. But the correlation is still sizable, and highly significant (0.58; $p < 0.001$), consistent with the presence of some fair types who put in non-minimal effort even in the final period.

through private contract offers play a bigger role when contracts are implicit. Firms are more selective in hiring specific workers. Consequently, the fraction of private contract offers is substantially higher in the IC treatment. Analyzing the hiring strategy at the firm level underlines these differences: 51% of firms in the IC treatment rehire the same worker in at least half of the market periods without interruption. The corresponding fraction of firms who does so in the C treatment is merely 3%. Overall, firms in the IC treatment are quite successful in eliciting effort through implicit performance incentives: while average effort is lower than under explicit contract enforcement, workers in the IC treatment choose an effort level of 8–10 in more than 50% of cases, and minimum effort is observed in only 12% of cases.

	C treatment	IC treatment	U-Test
1) Wages	22.80	36.60	$p < 0.01$
2) Rents offered by firms ($w - c(\hat{e})$)	5.74	21.12	$p < 0.01$
3) Realized worker rents ($w - c(e)$)	5.74	23.59	$p < 0.01$
4) Effort	9.65	6.97	$p < 0.01$
5) Wage-effort correlation	0.12	0.79	$p < 0.01$
6) Fraction of private contract offers	0.305	0.765	$p < 0.01$
7) Fraction of firms with employment relationships ≥ 9 market periods	0.029	0.514	$p < 0.01$
8) Fraction of possible vacancies offered	0.986	0.679	$p < 0.01$
9) Fraction of offered vacancies accepted	0.961	0.992	$p < 0.05$

Table 2: *Market characteristics. Means values across treatments; p-values (two-sided) are based on session averages.*

The difference in the number of contract offers between treatments deserves particular attention. We have hypothesized that the emergence of unemployment in the IC treatment might be a consequence of the necessity to pay higher rents for eliciting work effort. Since firms face decreasing returns to scale, they might not be able to pay sufficiently generous rents when employing two workers. Therefore, firms might prefer to ration job offers and employ

fewer workers than the maximum possible. The results on offered and accepted contracts (see the final two lines of Table 2) are in line with this prediction. In both treatments, the number of potential vacancies that firms could offer is 1260. While nearly 99% of these potential vacancies are actually offered in the C treatment (1242 out of 1260), firms in the IC treatment make only 68% (856 / 1260) of the possible contract offers. At the same time, workers in both treatments hardly ever reject an offered vacancy. The corresponding fraction of accepted vacancies is even slightly higher in the IC treatment.

Figure 2 sheds more light on firms' decision to ration jobs in the IC treatment. The dashed lines depict the fractions of firms who offer a second vacancy in a given period. The solid lines indicate the fractions of second vacancies filled, i.e., offers which are accepted by some worker. Paralleling the observation on the unemployment level, the fraction of firms who seek to hire two workers is initially similar in both treatments. After a few periods this fraction increases to more than 90% in the C treatment, but fewer and fewer firms decide to employ two workers in the IC treatment, in particular through periods 1-7. The solid lines confirm that unfilled vacancies, which would be an indication for voluntary unemployment, are rarely observed in either treatment. If anything, the fraction of rejected contract offers is somewhat higher in the C treatment.⁸ Again, this aggregate result is confirmed when we analyze the hiring strategy at the firm level. In the C treatment, a vast majority of firms employ two workers in nearly all periods: all firms employ two workers in at least 12 market periods, and 83% of firms hire two workers in 15 or more periods. The corresponding fraction of firms who do so in the IC treatment is only 17.1% and 11.4%, respectively.

The prevalence of job rationing, the high rents paid to workers and the low frequency of rejected job offers under contractual incompleteness indicate that unemployment in the IC treatment is involuntary.

Result 2: *When contracts are not explicitly enforceable, firms tend to ration job offers, build up long-term employment relationships, and pay generous rents to their worker. The resulting unemployment is involuntary.*

⁸cp. Line 9 of Table 2. Most of the rejected offers are contracts which offer zero or negative rents to the workers.

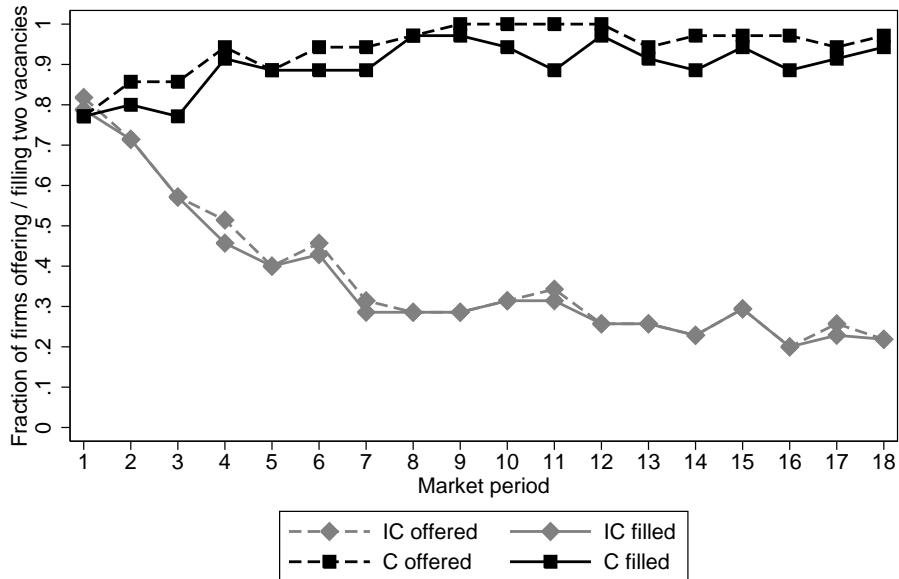


Figure 2: *Fraction of firms offering / filling two vacancies.*

4.2 Labor Market Segmentation as a Consequence of Contract Enforcement Problems

In the previous section we have seen that—when contracts are explicitly enforced—essentially all firms hire two workers and there is nearly full employment. Firms in the C treatment also pay wages close to the market clearing level and extract the major share of rents from production. Under contractual incompleteness, we have argued that it is attractive for firms to rely on long-term employment relationships, pay generous rents to workers and ration job offers. In the long run, we would thus expect all firms in the IC treatment to adopt a one-worker strategy.

A closer inspection of Figure 2, however, reveals that the latter prediction is only partly true. The figure confirms that in the C treatment nearly all firms employ two workers. Moreover, in the IC treatments there are strong dynamics towards a one-worker strategy in the first seven periods. In this phase, the decline in the fraction of firms who use a two-worker strategy is sizable and statistically significant (t-test, $p < 0.01$).⁹ However, the number of

⁹The reported test statistics come from regressing the fraction of two-worker firms in a given period and a given session of the IC treatment on a linear time trend for the early phase (periods 1–7) and late phase (periods 8–18), allowing for different slopes and intercepts for the different phases. The reported p-values

two-worker firms, i.e., firms who employ two workers in a given period, does not go all the way down to zero in later periods. Rather, after the initial strong decline in the number of two-worker firms a plateau seems to be reached where a relatively stable fraction of about 20–30% of firms continue to employ two workers. In this later phase (periods 8–18), the decline in the fraction of two-worker firms is essentially zero and not significant anymore (t-test, $p = 0.219$). The difference in dynamics between the two phases is statistically significant (t-test, $p < 0.05$). In other words, after a first phase where we observe strong dynamics towards adopting a one-worker strategy, two types of firms seem to coexist in later periods of the IC treatment. In contrast, all firms seem to employ the same hiring strategy in the C treatment.

A potential explanation for the observed dynamics in the IC treatment would be that two-worker firms were relatively less profitable than one-worker firms only in the early phase of the experiment, but that in the long run both strategies yielded similar profits for firms. A convergence of profits between the different firm types might explain why some firms choose to adopt a two-worker strategy even in the long run. In fact, there is a strong difference in firm profits during the first seven periods: in this phase, one-worker firms on average earn 37.21 points per period—12.55 points more than firms who employ two workers. The difference in firm profits is statistically significant (t-test, $p = 0.010$). In contrast, during the late phase of the experiment profits between one-worker firms and two-worker firms do not differ significantly anymore (the average difference is 1.42 points; t-test, $p = 0.847$). This convergence in firm profits is due to an increase of two-worker firms’ profitability. While profits of one-worker firms remain virtually unchanged (profits fall insignificantly by 0.48 points; t-test, $p = 0.843$), two-worker firms earn on average 10.65 points more than in the early phase of the experiment (t-test, $p = 0.036$).

Figure 3 (top panel) confirms the profit differences between one-worker firms and two-worker firms in early periods. The figure is based on parameter estimates for firm profits in the first seven periods of the IC treatment, depending on whether firms employ one or two workers in a given period. As regressors, we include the wage paid per worker and the squared value of the wage to account for possible non-linearities in the profit-wage relation.

are adjusted for clustering on the session level.

The figure confirms that firm profits tend to increase in wage payments up to a maximum after which it does not pay off for firms to further increase worker rents. Importantly, the figure shows that two-worker firms tend to earn substantially lower profits irrespective of the wage they paid to their workers. Consequently, the profit-maximizing strategy for firms during the early phase is to employ only one worker at a relatively generous wage.

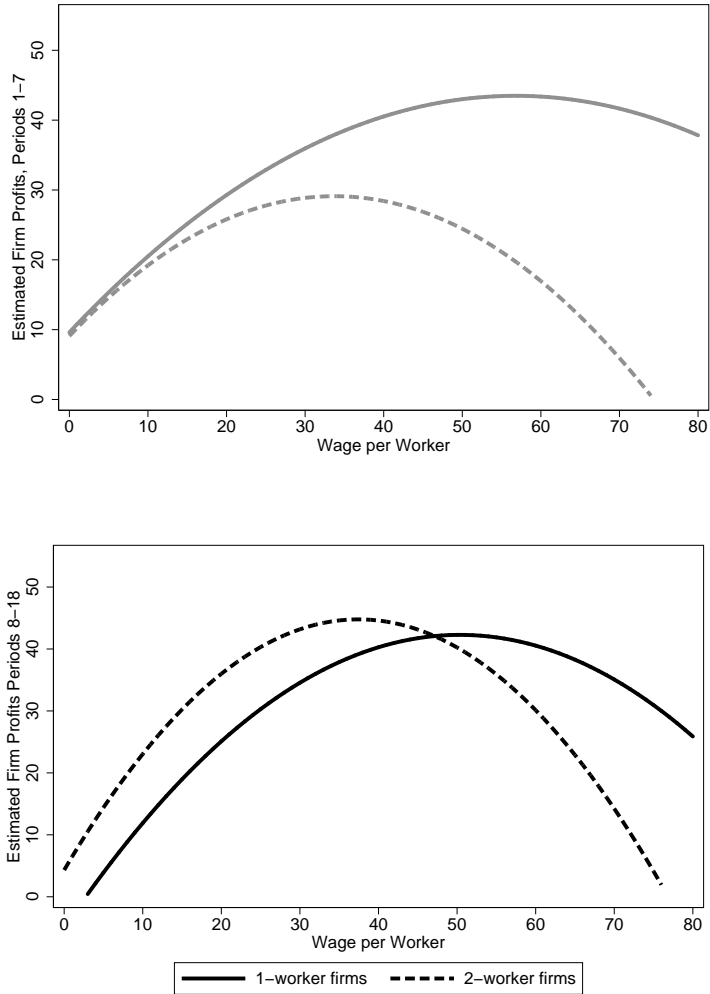


Figure 3: Predicted total profits of one-worker firms and two-worker firms in the IC treatment, depending on the wage paid per worker. Upper panel: periods 1-7, bottom panel: periods 8-18.

Turning to the *later phase* of the experiment (bottom panel of Figure 3), one-worker firms continue to be similarly profitable compared to early periods. However, profits for two-worker firms change dramatically. According to our estimates, two-worker firms earn

similar profits as one-worker firms in this phase when choosing the profit-maximizing wage. The fact that—in the long run—profits of one-worker firms and two-worker firms are similar suggests that the contract enforcement problems in the IC treatment cause a labor market segmentation. In the absence of explicit contract enforcement, two different firm strategies coexist in the same market: a majority of firms employ a one-worker strategy. At the same time, a second segment exists, consisting of firms who profitably use a different contracting strategy.

Indeed, considering the later market phase of the IC treatment, the two segments qualitatively differ in several important dimensions other than firm size. While both strategies yield similar profits for firms, workers in two-worker firms face much less favorable contract terms than their counterparts in one-worker firms. Wages in one-worker firms are on average 42.33 points—14.63 points higher than those for workers employed in two-worker firms. This difference indicates that one-worker firms pay higher rents than two-worker firms. Indeed, rents offered to workers in one-worker firms are almost 70% higher than those offered by firms who employ two workers. While workers in one-worker firms also work somewhat harder (the difference in effort levels is 1.43), the substantially higher wages in one-worker firms result in much higher realized rents for workers in those firms. On average, rents for workers employed in one-worker firms are 29.21 points whereas rents for workers in two-worker firms amount to 17.55 points per period. The two segments also seem to differ in the hiring and worker retention strategies used by firms: 90.04% of firm-worker matches in one-worker firms are initiated through private contract offers. This number is 12.5 percentage points higher than in two-worker firms. In addition, turnover rates differ substantially between the two different firm types: a worker’s likelihood of being dismissed after a given period is 23.78% in one-worker firms and 46.20% in two-worker firms.

Result 3: *Contract enforcement problems lead to a labor market segmentation. In the long run, two different types of firms coexist in the market with incomplete contracts. These earn similar profits, but differ qualitatively with respect to turnover and worker rents.*

4.3 Explaining Labor Market Segmentation

Our evidence so far indicates that contract enforcement problems cause both involuntary unemployment and a segmentation of the labor market. This raises the question what drives the segmentation of the labor market in the IC treatment. In particular, it is interesting to analyze why the profitability of two-worker firms increases in the long run and eventually converges to the level of one-worker firms (see Figure 3). Since firm profits in our setup are solely determined by wage payments and workers' effort provision, it has to be the case that in the late phase workers in two-worker firms exert higher levels of effort for a given wage compared to the early phase of the experiment. In fact, an estimation of wage-effort relations in two-worker firms during the early and late phase indicates that workers tend to provide significantly higher effort for a given wage during the late phase.¹⁰ The question is why this change occurs.

There are two plausible reasons for why we observe this pattern. First, it could be the case that workers in our experiment differ in their inherent willingness to provide effort for a certain wage. To operate successfully, firms who employ a two-worker strategy have to find workers who are willing to provide high effort for a lower wage than firms who employ only one worker. This selection process of "less demanding" workers into two-worker firms might take some time, and therefore the profitability of two-worker firms increases only once appropriate workers have been found. The second potential explanation for the increase in two-worker firms' profitability is that workers adapt their behavior over the course of the experiment. In other words, workers in two-worker firms might change their willingness to provide effort for a given wage in response to the observed market dynamics. The high level of unemployment and the prevalence of job rationing by one-worker firms in later periods of the IC treatment increases the cost of being unemployed. This is underlined by the fact that the job acquisition rate, i.e., the fraction of previously unemployed workers who find employment in a given period, drops from more than 70% in the first periods to about 25%

¹⁰This result is based on an OLS regression of effort on wage, a dummy variable for the late phase of the experiment, and an interaction term between the dummy and the wage. The sample is restricted to two-worker firms. The interaction term is significant ($p = 0.015$) and indicates a 27 percent increase in the marginal effect of wage on effort, relative to the early phase.

in the later phase of the IC treatment. In the long run, workers in two-worker firms might therefore *become* more willing to exert higher effort for a given wage in order to avoid job losses.

To disentangle the two explanations, we first analyze individual behavior of workers who are predominantly employed in two-worker firms during the late phase. That is, we concentrate on workers who work in a two-worker firm in at least 50% of their employment spells between period 8 and 18. If selection of workers who are willing to perform for low wages is the explanation for the increase in profits of two-worker firms, these workers should already exhibit high efforts for a given wage level in the early phase of the experiment. If, in contrast, labor market segmentation and the increased profitability of two-worker firms are due to market feedback mechanisms, we should observe a change in the behavior of those workers. More precisely, we should observe an increase in effort for a given wage between the early phase and the late phase of the experiment.

Figure 4 compares behavior of workers who are mostly employed in two-worker firms during the late phase to the behavior of *the same workers* earlier in the experiment. The figure shows a clear increase in effort provision between the early and the late phase for the workers. The increase in effort holds for the entire range of wage levels, and in particular for wages below 40. This is important since contracts involving such relatively low wages account for about 70% of contracts in two-worker firms and are, at the same time, the contracts that are particularly profitable for two-worker firms (see Figure 3).¹¹ This result strongly suggests that the increase in two-worker firms' profitability is attributable to a response of workers to the tighter labor market conditions rather than to selecting workers who have an inherently higher willingness to provide effort.¹²

We also find evidence for a link between market conditions, and the probability of shirk-

¹¹An OLS regression shows that the increase is statistically significant at the five percent level for several of the wage categories individually (for 30 to 39, $p = 0.035$; 40 to 49, $p = 0.018$), and jointly for wage intervals spanning the range 0 to 49 (F-test, $p < 0.001$).

¹²The limited role of worker selection is further underlined by the fact that, if anything, workers who are later on predominantly employed in two-worker firms tend to exert *lower* efforts for a given wage in the early phase of the experiment compared those workers who later are mostly employed in one-worker firms. See Figure 5 in the Appendix.

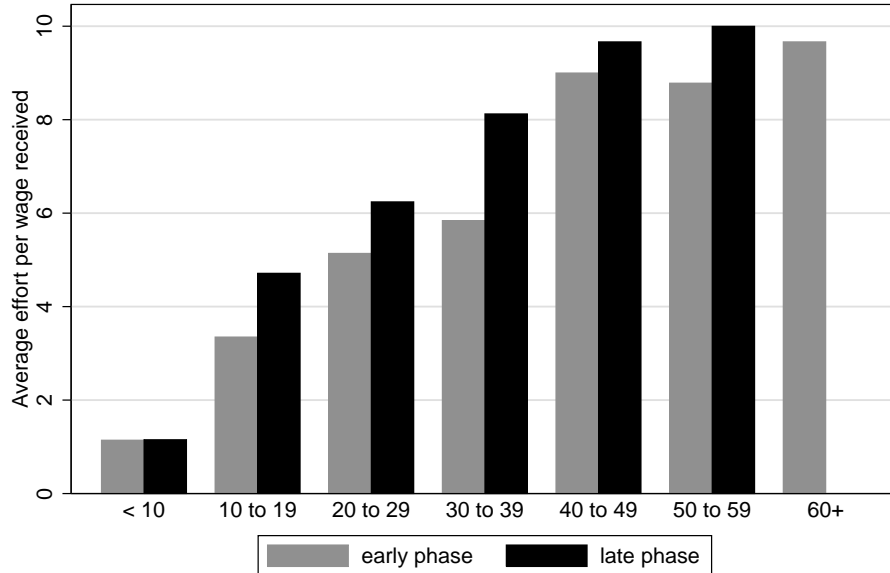


Figure 4: Average effort for a given wage. Values for workers who predominantly work in two-worker firms during the late phase of the experiment. To account for the strong decline in efforts and wages during the final periods (“endgame-effect”), the graph is based on values for periods 1–16 only.

ing, based on regression analysis. We estimate a Probit regression where the dependent variable equals 1 if the worker exerts less effort than contractually desired by the firm. Explanatory variables include the wage, the job acquisition probability in a given period (the share of unemployed workers who find a job in that period), and a dummy variable for whether the worker has personally experienced unemployment in the previous period. The estimates show that the falling job acquisition rate is associated with a fall in the probability of shirking ($p = 0.08$), and that the effect is sizable: According to the model the 50 percentage point drop in the job acquisition rate observed going from the early to late phase of the experiment implies a 26 percent drop in the probability of shirking. Interestingly, however, there is an additional effect of a worker having personally experienced unemployment: Controlling for the acquisition rate, and wages, having personally experienced unemployment leads to a statistically significant ($p = 0.013$) decrease in the probability of shirking, of about 20 percent. This suggests that, in our setup, personal experience of unemployment tends to discipline workers in addition to overall market tightness. It also implies that the

policy of frequently firing and hiring different workers, practiced by two-worker firms, serves to ensure that the firm has workers who are strongly motivated to work.

The feedback from market conditions to worker effort exertion thus helps to explain why we observe a segmentation in the market in the long run, with one-worker firms and two-worker firms coexisting. One-worker firms become more and more prevalent over time initially, increasing unemployment pressure, until the unemployment pressure is strong enough to make firms indifferent between adopting a one-worker or two-worker strategy. At this point there is no incentive for switching from one firm size to another, and segmentation emerges in terms of two different firm sizes.

Result 4: *Labor market segmentation and the increase in two-worker firms' profitability are attributable to an endogenous change in workers' behavior, which is associated with changing market conditions.*

5 Conclusions

In this paper, we have analyzed the relationship between contract enforcement and the emergence of involuntary unemployment. In an experimental labor market where trading parties can form long-term relationships, we compared a work environment where contracts were not third party enforceable to a situation where effort was verifiable and explicit contracts were feasible.

Our main findings are as follows: unemployment is much higher in the treatment without explicit contract enforcement. Importantly, unemployment in this treatment is involuntary, being caused by the firms' employment and contracting policy. Firms pay high wages but offer fewer vacancies than is possible and technologically efficient. This policy, however, succeeds in eliciting high efforts from the employed workers. When complete contracts can be written, wages are close to the market clearing level, firms do not ration jobs, and unemployment is basically absent. Thus, our results empirically establish a direct causal link between contract enforcement, the nature of employment relations, and the emergence of unemployment—a link which has since long been discussed in theory (Shapiro and Stiglitz 1984, MacLeod and Malcomson 1989, Akerlof and Yellen 1990, Sobel 2006).

Our findings also contribute to the literature on the efficiency-wage foundation of dual labor markets (Bulow and Summers 1986, Saint-Paul 1996). Although firms face identical technological constraints in our setup we observe a segmentation of the labor market, resulting from contractual incompleteness: when third party contract enforcement is not feasible, most firms ration the number of jobs and build up long-term employment relationships which are characterized by rent-sharing between firms and workers. The resulting increase in unemployment allows a minority of firms to operate successfully without rationing job offers. Tighter market conditions help these firms since stronger competition among workers makes hiring and dismissing workers easier and since workers shirk less. Interactions between these firms and their workers resemble those in secondary segments of the labor market (Doeringer and Piore 1971, Saint-Paul 1996). In contrast to one-worker firms, successful two-worker firms are characterized by lower rent payments and higher worker turnover.

References

- AKERLOF, G. A. (1982): “Labor Contracts as Partial Gift Exchange,” *Quarterly Journal of Economics*, 97, 543–569.
- AKERLOF, G. A., AND J. L. YELLEN (1990): “The Fair Wage-Effort Hypothesis and Unemployment,” *Quarterly Journal of Economics*, 105, 255–283.
- BAKER, G., R. GIBBONS, AND K. J. MURPHY (1992): “Relational Contracts and the Theory of the Firm,” *Quarterly Journal of Economics*, 117(1), 39–84.
- BEWLEY, T. (1999): *Why Wages Don't Fall During a Recession*. Harvard University Press, Cambridge, MA.
- BLANCHFLOWER, D. G., A. J. OSWALD, AND P. SANFREY (1996): “Wages, Profits, and Rent-Sharing,” *Quarterly Journal of Economics*, 111(1), 227–251.
- BROWN, M., A. FALK, AND E. FEHR (2004): “Relational Contracts and the Nature of Market Interactions,” *Econometrica*, 72(3), 747–780.
- BULOW, J. I., AND L. H. SUMMERS (1986): “A Theory of Dual Labor Markets with Application to Industrial Policy, Discrimination, and Keynesian Unemployment,” *Journal of Labor Economics*, 4(3), 376–414.
- BURKS, S. V., J. P. CARPENTER, AND E. VERHOOGEN (2007): “Fairness and Freight-Handlers: Local Labor Market Conditions and Wage Fairness Perceptions in a Trucking Firm,” *Industrial and Labor Relations Review*, 60, 477–498.
- CAMPBELL III, C. M., AND K. S. KAMLANI (1997): “The Reasons for Wage Rigidity: Evidence from a Survey of Firms,” *Quarterly Journal of Economics*, 112(3), 759–789.
- COHN, A., E. FEHR, AND L. GÖTTE (2008): “Fairness and Effort—Evidence from a Field Experiment,” *mimeo, University of Zurich*.
- DOERINGER, P. B., AND M. J. PIORE (1971): *Internal Labor Markets and Manpower Analysis*. Heath Lexington Books, Lexington, MA.

- FALK, A., AND J. J. HECKMAN (2009): “Lab Experiments Are a Major Source of Knowledge in the Social Sciences,” *Science*, 326(5952), 535–538.
- FEHR, E., AND A. FALK (1999): “Wage Rigidity in a Competitive Incomplete Contract Market,” *Journal of Political Economy*, 107(1), 106–134.
- FEHR, E., AND S. GÄCHTER (2000): “Fairness and Retaliation: The Economics of Reciprocity,” *Journal of Economic Perspectives*, 14, 159–181.
- FEHR, E., S. GÄCHTER, AND G. KIRCHSTEIGER (1997): “Reciprocity as a Contract Enforcement Device: Experimental Evidence,” *Econometrica*, 65, 833–860.
- FEHR, E., G. KIRCHSTEIGER, AND A. RIEDL (1996): “Involuntary Unemployment and Non-Compensating Wage Differentials in an Experimental Labour Market,” *Economic Journal*, 106, 106–121.
- FEHR, E., AND K. M. SCHMIDT (1999): “A Theory of Fairness, Competition and Cooperation,” *Quarterly Journal of Economics*, 114, 817–868.
- FISCHBACHER, U. (2007): “z-Tree: Zurich Toolbox for Ready-made Economic Experiments,” *Experimental Economics*, 10, 171–178.
- GREINER, B. (2003): “An Online Recruitment System for Economic Experiments,” in *Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht 63*, ed. by K. Kremer, and V. Macho, pp. 79–93. Ges. für Wiss. Datenverarbeitung, Göttingen.
- KRUEGER, A. B., AND L. H. SUMMERS (1988): “Efficiency Wages and the Inter-Industry Wage Structure,” *Econometrica*, 56(2), 259–293.
- KUBE, S., M. MARÉCHAL, AND C. PUPPE (2008): “The Currency of Reciprocity—Gift-Exchange in the Workplace,” *IEW Working Paper*, 377.
- MACLEOD, W. B., AND J. M. MALCOMSON (1989): “Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment,” *Econometrica*, 57(2), 447–480.
- (1998): “Motivation and Markets,” *American Economic Review*, 88(3), 388–411.

SAINT-PAUL, G. (1996): *Dual Labor Markets – A Macroeconomic Perspective*. MIT Press, Cambridge, MA, 1 edn.

SHAPIRO, C., AND J. E. STIGLITZ (1984): “Equilibrium Unemployment as a Worker Discipline Device,” *American Economic Review*, 74(3), 433–444.

SOBEL, J. (2006): “For Better or Forever: Formal versus Informal Enforcement,” *Journal of Labor Economics*, 24(2), 271–297.

6 Appendix

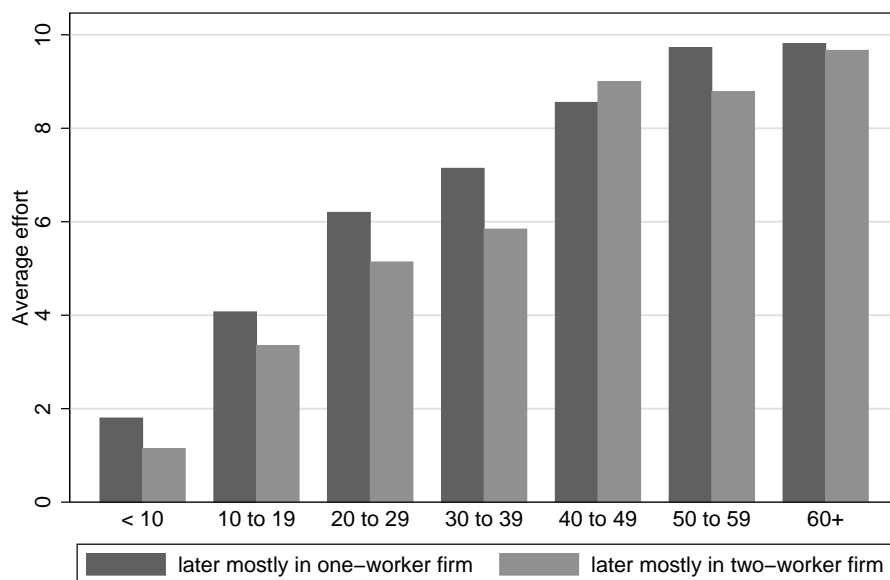


Figure 5: *Average effort for a given wage during periods 1–7. Workers who work predominantly in one-worker firms vs. workers who work predominantly in two-worker firms during the late phase (i.e., in more than 50% of their employment spells in periods 8–18).*