

THE COMPLEMENTARITY BETWEEN TRUST AND CONTRACT ENFORCEMENT*

Björn Bartling, Ernst Fehr, David Huffman and Nick Netzer

Agreements are rarely fully enforceable, exposing trading parties to the risk of exploitation. In such cases, trust—defined as the belief in the trustworthiness of others—may be key to realising gains from trade. This paper asks whether trust substitutes for or complements contract enforcement. While intuition may suggest that trust matters more when enforcement is weak, our experiment, which exogenously varies both trust and enforcement, shows that it can matter more when enforcement is strong. The mechanism is equilibrium selection: stronger enforcement allows more equilibria, and trust helps select efficient ones. This complementarity between trust and enforcement has important policy implications.

JEL codes: C91, D02, D91, E02

Incomplete and imperfectly enforceable agreements are ubiquitous in economic life. One reason for this are informational constraints, which make it impossible in most cases to govern all conceivable contingencies in a contract, and which can sometimes also prevent the verification of all enforcement-relevant information. For example, it may be difficult or impossible to obtain accurate information about past behaviour, especially in settings where trading partners are often strangers, or where it is costly to observe and record performance. This can preclude key strategies for enforcing agreements, such as conditioning contract terms on past performance. Even if informational constraints do not bind, other types of enforcement limitations can exist, such as lack of commitment power. For example, weak judicial systems can limit the ability of trading parties to legally commit to the specified contract terms.

Given these imperfections, contracting parties are exposed to the risk of being cheated, and they may only be willing to interact and realise the associated gains from trade if they trust that the other party will not take advantage of them. Consequently, it has been argued that trust is of fundamental importance for achieving economic efficiency (see, e.g., Banfield, 1958; Arrow, 1972; Coleman, 1990; Putnam, 1993; 2000; Fukuyama, 1995). The potential impact of trust on economic outcomes appears extensive, as it can affect individual-level economic interactions, organisational and market efficiency, and even economic development at the country level.

While there seems little doubt that trust is relevant when contracts are incomplete, this paper addresses a previously unstudied question: whether trust, defined in this paper as a belief about

* Corresponding author. Björn Bartling, Department of Economics, University of Zurich, Blümlisalpstrasse 10, 8006 Zurich, Switzerland. Email: bjoern.bartling@econ.uzh.ch

This paper was received on 1 November 2022 and accepted on 26 August 2025. The Editor was Steffen Huck.

The data and codes for this paper are available on the Journal repository. They were checked for their ability to reproduce the results presented in the paper. The replication package for this paper is available at the following address: <https://doi.org/10.5281/zenodo.16930021>.

The experiments reported in this paper were exempt from ethics board approval, in accordance with the regulations of the Human Subjects Committee of the Faculty of Economics, Business Administration, and Information Technology at the University of Zurich, which were in place at the time. The additional TRANSPARENCY-LT treatment, conducted subsequently as part of the same project during the revision process, did not undergo separate ethics approval. Earlier versions of this paper were circulated under the titles ‘The Causal Effect of Trust’ and ‘The Complementary Nature of Trust and Contract Enforcement’.

the trustworthiness of others, is a *substitute* for or a *complement* to the degree of contract enforcement opportunities. The ability to enforce contracts is arguably always imperfect, but the degree of imperfection varies substantially across time and space, due to factors such as technology, geography and institutions. Given this variation, one intuition might be that high trust is particularly important in environments with limited enforcement opportunities, as it can compensate for these limitations. In other words, trust and the strength of contract enforcement might be viewed as substitutes. Indeed, a perfectly enforceable contract is clearly a substitute for trust as there is no role for trust when contracts are complete.

In this paper, however, we demonstrate that trust and strength of contract enforcement can also be complements. We argue that the mechanism responsible for a complementary relation between trust and contract enforcement is equilibrium selection. Indeed, unless variation in trust, which is a variation in initial beliefs, is selecting different equilibria, it is difficult to understand how trust can have stable, lasting effects. If expanding contract enforcement opportunities simultaneously gives rise to a wider range of equilibria then trust may be more important when contract enforcement is stronger, though not perfect. In such a scenario, initial high trust beliefs could select efficient equilibria confirming these beliefs, while initial low trust beliefs could select inefficient equilibria, again confirming the beliefs. Conversely, in weak contracting environments with only a narrow range of (presumably inefficient) equilibria, trust may have little impact.

Understanding whether trust and contract enforcement are substitutes or complements is important due to differing policy implications. If they are substitutes, policies aimed at enhancing economic performance might be effective if they focused solely on improving formal institutions, such as the judicial system, to enable better contract enforcement—even if levels of trust remain low. Likewise, policies aimed at moving a society out of a low-trust trap, like public awareness campaigns promoting codes of conduct or advertising role models of trustful business relations, could be effective—even if formal institutions remain weak and ensure only very imperfect contract enforcement. However, if trust and contract enforcement are complements, the above policies might be ineffective if pursued in isolation. Effective policies would then need to simultaneously improve formal institutions, thereby allowing a wider range of more efficient equilibria, *and* raise trust levels, fostering the selection of efficient equilibria.

In this paper we report the results of controlled experiments showing that an independent improvement in contract enforcement opportunities at low levels of trust generates no or only small increases in gains from trade. In contrast, improvements in contract enforcement opportunities cause large increases in gains from trade if trust levels are high. Likewise, our data show that an increase in trust leads to no improvement in gains from trade if the environment is unfavourable for contract enforcement, but to large increases in gains from trade when the opportunities for contract enforcement are enriched. Our results are based on the exogenous variation of trust and the availability of different levers for contract enforcement in a laboratory experiment involving principals and agents facing profitable trading opportunities. The key advantage of this approach is that it allows for a clean separation of the effects of trust and strength of contract enforcement and their interaction on the realised gains from trade.

What economic and psychological mechanisms drive the complementarity between trust and contract enforcement? To provide deeper intuitions, we need to provide more detail about the experiment. The experiment involves principal-agent interactions where principals make contract offers in an experimental market by promising to pay wages and requesting effort levels from the agents, while agents choose their actual effort levels after accepting an offer. The gains from trade increase with effort, but there is a conflict of interest as higher effort benefits the principal while

being costly for the agent. In all treatments, there is a baseline degree of imperfection in contract enforcement captured by allowing agents to choose any effort level, regardless of contract terms. Subjects interact in markets of seven principals and ten agents over fifteen periods. In a given period, a match occurs if an agent accepts a principal's offer.

We implemented variation in the favourability of the environment for contract enforcement by altering the availability of two enforcement levers. One lever is the ability to condition contract terms on enforcement-relevant information, which we implement as the ability of principals to keep track of the past performances of the agents they have met. The other lever is commitment power, which we implement as the principals' ability to make the specified wage in the offered contract legally binding.

We consider three contracting environments, which progressively enrich the possibilities for contract enforcement.¹ In our *weak contracting environment*, neither lever is available. The principal can pay any wage to the agent, irrespective of the wage that was promised in the contract, so there is no commitment.² The principal and the agent simultaneously choose the actual wage and actual effort after the agent accepted the contract. In addition, the parties face an informational constraint that prevents them from making contracts contingent on signals of past behaviour. We implement this constraint by re-randomising the identification numbers of principals and agents across periods such that interactions remain one-shot. In our *medium contracting environment*, we improve contract enforcement by adding the lever of commitment: the principals' wage promises are legally enforceable, i.e., the principal must pay the promised wage. Our *strong contracting environment* adds, on top, the lever of being able to condition contract terms on past performance by keeping identification numbers constant across rounds. This allows principals to make their contract offers contingent on signals about a specific agents' previous effort choices.

We implement (and verify) exogenous variation in principals' initial trust about the agents' trustworthiness. We achieve this by showing the principals, before the start of the experiment, examples of real historical effort choices in experimental sessions in which the agents either exhibit trustworthy behaviour (shown in our high-trust treatments) or untrustworthy behaviour (shown in our low-trust treatments). This design allows us to study a possible causal effect of trust by comparing high-trust and low-trust treatments.

How do we explain our finding that the impact of trust depends in a complementary manner on strength of contract enforcement, and vice versa? Our investigation of mechanisms suggests three key ingredients. (1) An important underlying mechanism for trustworthy behaviour is that some agents have a preference for reciprocity, rewarding high wages with high effort even in one-shot interactions, though this tendency is heterogeneous, with a sub-population of agents being selfish. (2) Commitment power is crucial for principals to elicit high effort levels from reciprocal agents and distinguish them from selfish agents; if wage promises are not credible, reciprocal types will not reveal themselves by choosing high effort. (3) Information about past

¹ As an extension, we also consider a fourth contracting environment that further strengthens contract enforcement to explicitly demonstrate that sufficiently strong contract enforcement can substitute for trust, which is trivially true in the limit where contracts are complete.

² This contracting environment may reflect a weak or inefficient judicial system or rely on informal agreements, often based on verbal promises. Such conditions are frequently observed in developing countries, where a large proportion of workers, sometimes even the majority, are employed in the informal sector, often on a casual basis (Banerjee and Duflo, 2007; Djankov *et al.*, 2008; La Porta and Shleifer, 2014; McCaig and Pavcnik, 2015). In this sector, written contracts are typically absent, the judicial system is highly inefficient or non-existent, and contract enforcement is consequently very weak.

performance matters for the ability of principals to benefit from identifying reciprocal types. This allows them to selectively condition high wages on signals of past trustworthy behaviour and engage in reciprocal relationships with high gains from trade.

We argue, and verify empirically, that these ingredients can give rise to the observed complementarity. Intuitively, if both commitment power and information are available, high trust can motivate principals to try a strategy of paying high wages and screening for reciprocal types. In contrast, low trust can cause principals to refrain from this strategy and pay only low wages. However, if commitment power or information is missing, even high trust among principals results in inefficient outcomes as there is no equilibrium allowing principals to both identify and benefit from reciprocal interactions with reciprocal types.

We begin our analysis of mechanisms with the weak contracting environment. In this setting, principals typically promise to pay high wages, but these promises are not contractually enforceable. As a result, principals rarely fulfil their promises and actually pay very low wages. Agents quickly realise that they cannot rely on the principals' promises, leading even reciprocal agents to show little willingness to respond to high offered wages with high effort. Principals have little reason to keep their promises since they experience no return from doing so. In other words, the lack of legal enforcement of wage promises undermines agents' reciprocal behaviour and generates a 'low-wage-low-effort' equilibrium.

In the final part of the paper, we present a theoretical model that captures key features of our experimental game and the heterogeneity in agent types in a simplified way. The model explains the empirical regularities in our weak contracting environment. It shows that the 'low-wage-low-effort' equilibrium is unique and predicts, in particular, that an exogenous shock to the principals' beliefs about the agents' trustworthiness has no effects on wages, effort and gains from trade, which aligns with our observations in this contracting environment.

The legal enforcement of the principals' wage promises in our medium contract environment constitutes a major improvement in contract enforcement by allowing the principals to credibly commit to high wages. Consequently, the agents understand that a high wage promise indeed offers them a generous share of the surplus, prompting reciprocal agents to respond to higher wages with higher effort levels. Our high-trust manipulation fosters optimistic beliefs about agents' trustworthiness, leading principals to expect that agents will reciprocate high wages with high effort. Conversely, in our low-trust environment, principals anticipate a weaker reciprocal response. Therefore, principals are inclined to offer higher wages in the high-trust environment compared to the low-trust environment, as confirmed by the data. Reciprocally motivated agents then respond with higher effort levels in the high-trust environment compared to the low-trust environment, which explains the positive average trust effect in the medium contracting environment.

Is the trust effect on the gains from trade in the medium contracting environment stable, i.e., an equilibrium phenomenon? Or is it due to a transitory effect of changing the principals' initial beliefs about the agents' trustworthiness? To answer this question, we need to examine whether principals benefited on average from paying high wages in the high-trust environment. It turns out that they did not; the agents' average effort increase in response to a wage increase is insufficient to render the wage increase profitable on average. This modest average response reflects agent heterogeneity, with some agents exhibiting a strong response, while others choose minimal effort regardless of the wage. Consequently, paying higher wages in the high-trust environment is not profitable on average, but principals learn this only slowly over time. This learning process

is indicated by the fact that wage offers in the high-trust environment decrease over time and eventually approach the levels in the low-trust environment.

Our theoretical model rationalises the transitory nature of the increase in the gains from trade in the medium contracting environment. The model shows that if the agents' effort responses are too small on average (e.g., because the share of reciprocally motivated agents is too small), there is still a unique 'low-wage–low-effort' equilibrium. However, the model also predicts that initially false (i.e., too optimistic) beliefs of the principals about the agents' trustworthiness induce the principals to initially make too high wage offers.

Finally, we show that there is a large and stable trust effect on agent effort levels, and thus the gains from trade, in our strong contracting environment. In this environment, principals can, not only make credible wage promises, but also condition their current contract offers on the agents' past performance signals. Principals can exploit this by making no offer, or an offer with a lower wage, to agents with low previous performance signals, while targeting high wage offers to agents with high previous performance signals. Empirically, principals in the high-trust environment indeed screen agents in this way and target their high wage offers to agents who previously signalled their trustworthiness. Consequently, the wage-effort relation is steeper in the strong contracting environment compared to the medium contracting environment. This implies that high wages can be profitable in the high-trust environment. In contrast, in the low-trust environment, the principals believe that the wage-effort relation is relatively flat and, therefore, make only low wage offers from the outset, choosing not to screen for reciprocal agents who respond to high wages with high effort levels. These wage differences between the high- and the low-trust environments are large and stable over time, inducing large and stable effort differences based on the agents' reciprocal effort responses.³

Our theoretical model rationalises these findings for the strong contracting environment. Specifically, the model shows that both a high-trust screening equilibrium and a low-trust pooling equilibrium coexist in this environment. The initial variation in the principals' trust appears to cause stable variation in the realised gains from trade by selecting between these equilibria.

Our paper makes several contributions to the literature. First, it documents experimentally that the effects of improvements in contract enforcement on gains from trade can be trust dependent. To our knowledge, this is a novel empirical finding that may generally be interesting for the economics of contracts and institutions (e.g., North, 1991; Bolton and Dewatripont, 2005) and, in particular, for behavioural contract theory that examines the effects of non-standard motives and social norms on the functioning of contracts and incentives (e.g., Ellingsen and Johannesson, 2005; 2008; Sliwka, 2007; Hart and Moore, 2008; Hart, 2009; Hart and Holmström, 2010; Herweg and Schmidt, 2015; Bierbrauer and Netzer, 2016; Danilov and Sliwka, 2017; Sliwka and Werner, 2017). Our findings on the role of trust in our strong contracting environment are also of interest for the literature on relational contracting (e.g., Gibbons, 1998; MacLeod and Malcolmson, 1998; Baker *et al.*, 2002; MacLeod, 2007; Gibbons *et al.*, 2023).

Second, our paper clarifies the conditions under which we can expect a causal effect of trust on gains from trade, and it shows the important role of the contracting environment in the transmission of initial trust differences on wages, effort levels and gains from trade. Our paper

³ However, we observe a pronounced endgame effect in the high-trust environment within the strong contracting context, which diminishes the differences in effort. Endgame effects, which we discuss in more detail later in the paper, are an artefact commonly found in many experimental gift-exchange games where the parties have opportunities for repeated interactions because, due to fixed identification numbers, they can identify each other and deliberately trade with the same previous partners.

thus contributes to the debate on the effect of trust on economic outcomes (e.g., Knack and Keefer, 1997; La Porta *et al.*, 1997; Guiso *et al.*, 2009; Algan and Cahuc, 2010; Bloom *et al.*, 2012) by clarifying when we can expect no, only transitory or stable effects of changes in trust.⁴

Third, by empirically showing conditions for the emergence of a stable and efficient reciprocal principal-agent interaction, our paper is also related to the literature on reciprocal gift exchange and trust using laboratory experiments (Fehr *et al.*, 1993; Berg *et al.*, 1995; Brown *et al.*, 2004; Charness, 2004). Previous papers in the gift-exchange literature, as well as more recent papers on the counterproductive effects of sanctions and other measures that constrain shirking by agents, suggest that trust might be self-confirming (e.g., Bohnet *et al.*, 2001; Bohnet and Huck, 2004; Falk and Kosfeld, 2006; Bartling *et al.*, 2012). However, this literature neither addresses the interaction of exogenous variations in trust and contract enforcement, nor identifies the conditions or mechanisms under which we should expect trust to affect trading efficiency.

Finally, our paper presents a simple theoretical model that captures the main empirical regularities observed in our experiment. The goal of the model is not to provide an exact image of the experimental game, but rather to provide a framework that facilitates a coherent interpretation of the data. The model rationalises, in particular, why an exogenous increase in trust has no effect on gains from trade in the weak contracting environment, only a transitory effect in the medium contracting environment, and a stable effect in the strong contracting environment.

The theoretical literature has shown before that different levels of trust can arise in a given economic environment due to multiple equilibria (e.g., Tabellini, 2008; Aghion *et al.*, 2010) or multiple stable long-run outcomes of dynamic learning processes (e.g., Bower *et al.*, 1996). To the best of our knowledge, our experiment provides a first explicit test of the general idea that trust can play a role due to multiple equilibria: the empirical result that an exogenous increase in trust has no stable effect in a unique equilibrium environment, but leads to stable effects in a multiple equilibrium environment demonstrates this point. Our theoretical model, however, differs from the existing literature in two important ways. First, we follow a standard game-theoretic approach with fixed preferences, while Tabellini (2008) and Aghion *et al.* (2010) studied behaviour that is transmitted from generation to generation and co-evolves slowly with external institutions. Our theoretical and empirical results show that trust is malleable rather quickly and can have immediate and stable causal effects even with fixed preferences and institutions. Second, in models like Bower *et al.* (1996), where agents learn about a given population state, long-run levels of trust and economic efficiency cannot be manipulated by interventions that select between different equilibria. In contrast, we show that selecting the right equilibrium is a crucial factor in the design of organisations and mechanisms. This idea has also recently gained prominence in organisational economics where it has been argued that a deeper understanding of the forces that enable organisations to ‘build’ a more efficient equilibrium is key in understanding why some organisations persistently perform better than others (Gibbons and Henderson, 2012; Gibbons, 2020; Gibbons *et al.*, 2023).⁵ Third and perhaps most important, our paper is the first to show that trust and contract enforcement can be complements because the strength of contract enforcement determines whether multiple equilibria exist.

⁴ Our paper varies trust exogenously and examines the consequences of trust. There is also a literature that studies the individual and collective determinants of trust (e.g., Alesina and La Ferrara, 2005). For a review of the literature on the determinants of trust, see Fehr (2009).

⁵ In this context, the complementarity between trust and contract enforcement (i.e., incentives) is also important. Our findings suggest that to reap the available gains from trade it sometimes needs a change in incentives *and* a change in trust.

The remainder of the paper is organised as follows. Section 1 explains our experimental design and includes a manipulation check showing the effectiveness of our exogenous variation of trust. Section 2 presents our main empirical finding on the complementarity between trust and contractual enforcement. Section 3 discusses behavioural mechanisms behind our main empirical finding by analysing in detail how differences in the contractual environment shape the behaviour of principals and agents. Section 4 presents our theoretical analysis of the principal-agent game. Section 5 reports the results of a fourth contracting environment, demonstrating that contract enforcement can be sufficiently strong to serve as a substitute for trust. Section 6 concludes.

1. Experimental Design

We study the impact of an exogenous variation in principals' beliefs about the trustworthiness of agents on wages, effort and gains from trade. To study the interaction between exogenous changes in principals' trust and the contract enforcement environment, we also vary the degree to which parties can enforce contracts. We adopt a typical principal-agent framework where a higher effort level by the agent increases the principal's expected payoff, but providing higher effort is more costly for the agent. Principals and agents interact in an experimental market and we allow for fifteen market periods, so that we can study how wages, effort and gains from trade evolve over time. This feature allows studying whether exogenous changes in trust or contract enforcement have stable or only transitory effects.

The framework follows closely that of Brown *et al.* (2004) with the main difference that, in our design, principals cannot directly observe effort levels, but they receive an informative stochastic signal about the agents' effort choices, and higher effort levels are associated with an increase in the probability of observing a high signal. In many types of economic interactions it is not possible to precisely identify whether effort or (bad) luck is responsible for the observed output. The effort signal is observable by the principal and the agent, but it is not verifiable by third parties, and thus not directly contractible. Contracts are therefore necessarily incomplete and the effort choice of an agent cannot be legally enforced. The principal's belief that an agent is trustworthy may then be relevant for the principal's willingness to enter into a transaction with an agent and for the contract terms the principal offers. We define agents to be trustworthy when they are willing to reciprocate a high wage offer with a high effort choice although high effort reduces their material payoff. Untrustworthy agents, by contrast, always choose low effort levels irrespective of the offered wages.

Our treatments vary the degree to which the parties can enforce their agreements. In all treatments, the principal proposes a contract that offers a wage and requests an effort level from the agent. In our weak contracting enforcement environment (WEAK), however, neither the offered wage nor the requested effort level is legally enforceable. This environment thus represents a situation with weak legal institutions. Furthermore, the identities of principals and agents are not observable, and thus contract terms cannot be conditioned on the agent's past performance signal. In our medium contracting enforcement environment (MEDIUM) we increase the scope for contract enforcement by making the principals' wage offers legally binding. However, the agent is still free to choose any effort level and it is still not possible to make contracts contingent on past performance signals. In our strong contracting enforcement environment (STRONG), principals' wage offers are again legally binding and agents are still free to choose any effort, but the subjects now have fixed identification numbers over the course of the experiment. Principals can therefore target their offers to specific agents contingent on their past performance signals,

Table 1. *Agents' Cost Function.*

Effort	1	2	3	4	5	6	7	8	9
Cost	0	1	2	4	6	8	10	12	15

which is a further expansion of the set of contractible contingencies relative to the other treatments. Performance signals from past periods are only observed by the respective principals with whom the agent interacted in these periods.⁶

1.1. *Stage Game Payoffs*

If a principal and an agent agree to trade then the principal pays a wage $w \in \{1, \dots, 100\}$ to the agent and the agent chooses an effort level $e \in \{1, \dots, 9\}$. The agent's effort choice stochastically determines the value of the interaction for the principal. There are only two possible value levels, 100 and 10. The probability that the principal receives the high value is given by $e/10$, while the principal receives the low value with probability $1 - e/10$. The (expected) material payoffs of principals and agents are thus given by

$$E[\Pi^{\text{principal}}] = \begin{cases} 100 \cdot \frac{e}{10} + 10 \cdot \left(1 - \frac{e}{10}\right) - w & \text{if principal and agent interact,} \\ 0 & \text{otherwise,} \end{cases}$$
$$\Pi^{\text{agent}} = \begin{cases} w - c(e) & \text{if principal and agent interact,} \\ 5 & \text{otherwise,} \end{cases}$$

where $c(e)$ denotes the agent's cost of providing effort. The outside option of an agent who does not interact with a principal is 5. Table 1 shows the cost function $c(e)$. The effort cost function is strictly increasing, with weakly increasing marginal costs. Given that the marginal cost of effort does not exceed 3, while the marginal expected benefit remains constant at 9, the efficient effort level is $e = 9$. The expected gains from trade are defined as the sum of the principal's and agent's (expected) payoffs minus the agent's opportunity cost of 5.

1.2. *Contracting Environments*

Principals can initiate trades by offering contracts to the agents. Agents can choose among the available contract offers, but they cannot make offers to principals. There are fifteen trading periods. Each period, a principal can interact with at most one agent, and an agent can accept at most one contract offer. A period has two stages. In stage one, the principals make contract offers and agents decide whether or not to accept a contract. If a principal and an agent conclude a contract, they enter stage two, where the principal pays a wage and the agent chooses an effort level.

A contract offer includes a wage offer $w \in \{1, \dots, 100\}$, a requested effort level $\tilde{e} \in \{1, \dots, 9\}$ and the principal's identification number (ID). The wage offer w is third-party verifiable, and thus contractible in treatments MEDIUM and STRONG, but not in treatment WEAK. The requested effort level $\tilde{e}$ is not binding. The principal can observe the realised value, but not the agent's

⁶ A fourth contracting environment, discussed as an extension in Section 5 below, allows principals to observe, not only the performance signals of the agents they directly interacted with, but also the performance signals of all other agents in the market. We thank two anonymous referees for suggesting this additional treatment.

actual effort choice. Indeed, actual effort levels e are never third-party verifiable, which rules out that requested effort levels are legally enforceable in any of our contracting environments.

There are two types of offers, public and private. In private offers, a principal specifies the ID of the agent they wish to trade with, and only this agent is notified about the contract offer. In contrast, public offers are visible to all agents as well as the other principals, allowing any agent the opportunity to accept them. A principal may issue an unlimited number of both private and public offers in a given period. However, once an agent accepts an offer, the principal is matched with that agent, learns their ID (which is news in the case of a public offer) and all remaining outstanding offers from that principal are withdrawn from the market.⁷ At the start of each period, no contract exists, and no offers have been made. The market consistently consists of ten agents and seven principals, resulting in an excess supply of three agents.

At the conclusion of each period, every subject receives information about their own payoff and a reminder of the contract $(w, \tilde{e})$ they agreed to, along with the ID of their trading partner. Agents are also informed about the payoff of their respective principal. Principals are not informed about the payoff of their respective agent, because a principal does not observe the agent's effort choice, and thus the cost of providing this effort level. Subjects record this information on a form supplied alongside the experimental instructions. This procedure ensures that they can consistently reference their own trading history.

1.2.1. *Contracting environment WEAK*

An agent can choose any actual effort level $e \in \{1, \dots, 9\}$ after having accepted a contract offer in our contracting environment WEAK, irrespective of the requested effort level $\tilde{e}$. Likewise, the principal can pay any wage to the agent, irrespective of the offered wage. Actual wages and effort levels are chosen simultaneously at the second stage of a period. Moreover, the subjects' IDs are randomly reshuffled in each of the fifteen periods of the experiment in contracting environment WEAK. Random IDs preclude the principals from conditioning future contract offers on past performance signals. Neither of the two contracting parties thus faces legal or economic incentives to stick to the terms of the contract in contracting environment WEAK, but intrinsic motivations could still induce them to honour their mutual promises.

1.2.2. *Contracting environment MEDIUM*

A principal is obliged to pay the offered wage if an agent accepts his or her contract offer in our contracting environment MEDIUM. An agent, however, can still choose any actual effort level $e \in \{1, \dots, 9\}$, irrespective of the requested effort level $\tilde{e}$. Principals must thus stick to the terms of the contract in contracting environment MEDIUM, while agents face no legal or economic incentives to provide the requested effort level. Because IDs are randomly shuffled in every period, as in contracting environment WEAK, principals cannot condition hiring and contract terms on signals about the past performance of specific agents.

1.2.3. *Contracting environment STRONG*

Contract enforcement is strengthened further in our contracting environment STRONG. The principal is obliged to pay the offered wage w if an agent accepts the contract, as in contracting environment MEDIUM. Moreover, while agents can still choose any actual effort level $e \in \{1, \dots, 9\}$, irrespective of the requested effort level $\tilde{e}$, IDs of all players are fixed in contracting

⁷ To ensure that principals do not extend private offers to agents who have already entered into a contract, they are continuously informed about which agents are still available in the market.

environment STRONG. This feature provides principals with the opportunity to condition their contract offers on the identity of a specific agent and this agent's past performance signals. This provides principals with the ability to screen agents and selectively target high wages to those who have high past performance signals.⁸

1.2.4. *Discussion of contracting environments*

One may ask which real-world environments correspond to our three contracting environments. In our view, many developing countries exhibit characteristics of the contracting environments WEAK and MEDIUM, as a substantial portion of economic activity occurs in the informal sector. In this sector, written contracts are rare, legal enforcement is weak or non-existent, population registries that facilitate the spatial identification of citizens, and reputational institutions, such as credit registries, are also lacking (Banerjee and Duflo, 2007; Djankov *et al.*, 2008; La Porta and Shleifer, 2014; McCaig and Pavcnik, 2015). Consequently, contract enforcement is often fragile.

However, it is important to recognise that even in relatively advanced economies, contract enforcement can sometimes be very inefficient. For example, in the United States, the absence of a population registry allows individuals to obscure their location, making it difficult to hold them accountable for contract breaches. Moreover, large segments of the population may not be covered by credit registries, and the legal system can also be very inefficient, as highlighted by Djankov *et al.* (2008).⁹

1.3. *Inducing Variation in Principals' Trust*

To induce exogenous variation in the principals' initial trust levels, we randomly assigned them to two different information conditions. In the high-trust treatments, the principals were informed about a historical example in which agents behaved in a trustworthy manner; in the low-trust treatments, they were shown an example in which agents displayed a low level of trustworthiness. More specifically, for our high-trust treatments, we selected the market from Brown *et al.* (2004) that had the steepest wage-effort relation, and for the low-trust treatments, we selected the market with the flattest wage-effort relation.

The example was provided at the end of the experimental instructions. Subjects were informed that the information provided was an 'example', and that it showed how effort is related to wage levels 'in a past session'. Subjects were told that the information in the example was something that they 'could use in their decisions today'. The description of the source of the example was completely truthful, but deliberately vague, and we did not claim that the information provided about a single past session was representative.

Figure 1 shows how we presented the examples in the experimental instructions. The top row was shown to the principals in the high-trust treatments, the bottom row to the principals in the low-trust treatments. On the left, the wage-effort relation is shown. The figure shows the average effort provided by the agents in the example for each of the given bins of offered wages. On the right, we show how this wage-effort relation translates into a wage-payoff relation, given the principals' payoff function in our experiment. The high-trustworthiness example involved

⁸ Wibral (2015) studied economic environments in which individuals can change their identities. The idea of inducing one-shot play by re-randomising IDs every period and enabling long-run relationships by (i) fixing IDs throughout the experiment and (ii) allowing for private offers to specific agents is taken from Brown *et al.* (2004).

⁹ Djankov *et al.* (2008) reported, for instance, that, for every dollar of credit issued in Germany, a lender recovers, on average, only 57 cents through the legal system when attempting to enforce repayment from an unwilling borrower. The corresponding recovery rates for Switzerland and Italy are 60.4 cents and 45.3 cents, respectively.

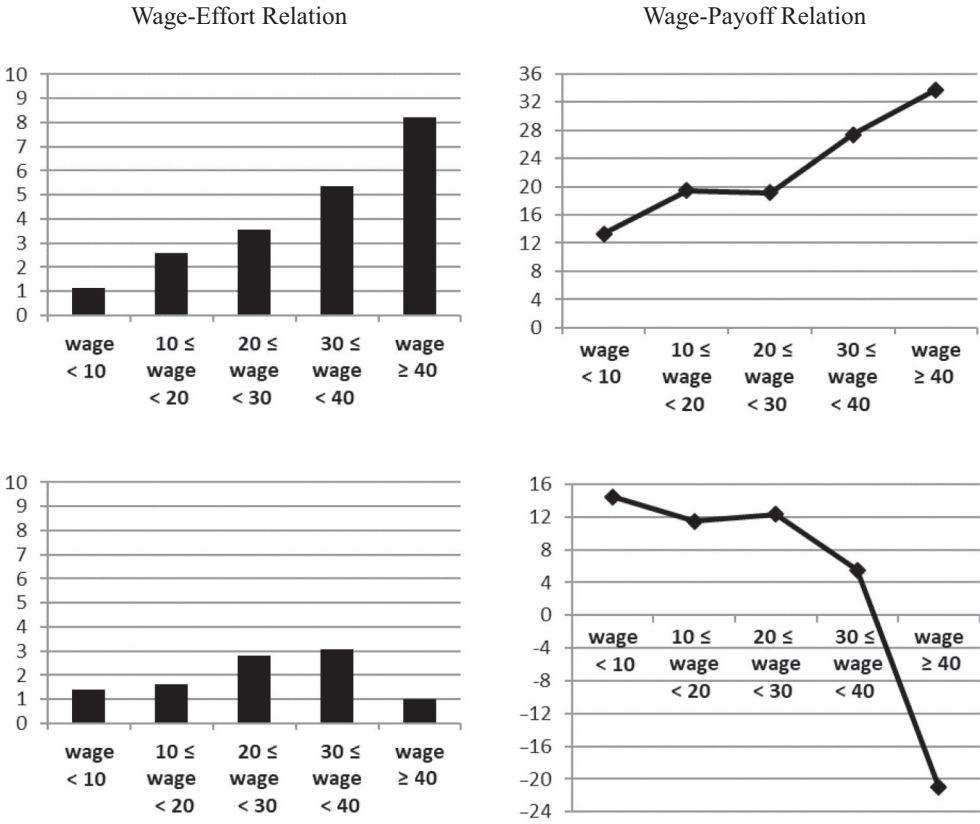


Fig. 1. *The High- and Low-Trustworthiness Examples Shown to Principals.*

Notes: The top row shows the example provided to the principals in the high-trust treatments; the bottom row shows the example provided in the low-trust treatments. The wage-effort relation is shown on the left and the corresponding expected wage-payoff relation on the right.

agents being trustworthy, in that those who are paid high wages also exert high effort levels. In the low-trustworthiness example, agents were untrustworthy; they provided rather low effort for all wage levels.

Note that the examples contain no information about the historical frequency of wage choices by principals. They do indicate the range of wages that was used, but this was identical across the high- and low-trustworthiness examples. This is deliberate because we wanted to rule out the possibility that the examples influence behaviour by conveying information about historical behaviour of principals. Rather, the differential information content across examples is solely about the trustworthiness of agents. Any impact should thus come through the beliefs of principals about trustworthiness. We examine in Section 1.5 below the extent to which our trust manipulation was effective in the sense that it differentially affected principals’ beliefs.

Note also that a given trust example holds constant the information given to principals about how agents respond to wages, regardless of the different contract enforcement environments. This reflects our goal to implement the same initial beliefs, to compare how the same level of trust

Table 2. *Treatment Overview.*

Treatment	Principals' wage offers	ID numbers	Trustworthiness example	Principals' information on agents' performance	# Periods	# Markets
WEAK-LT	Non-binding	Random	Low	Own agents only	15	5
WEAK-HT	Non-binding	Random	High	Own agents only	15	5
MEDIUM-LT	Binding	Random	Low	Own agents only	15	5
MEDIUM-HT	Binding	Random	High	Own agents only	15	5
STRONG-LT	Binding	Fixed	Low	Own agents only	15	5
STRONG-HT	Binding	Fixed	High	Own agents only	15	5
STRONG-HT-Long	Binding	Fixed	High	Own agents only	25	2
TRANSPARENCY-LT	Binding	Fixed	Low	All agents	15	5

may have different ultimate effects depending on the environment. In our manipulation check in Section 1.5 below we verified that a given trust example successfully implemented the same initial beliefs across contracting environments. Initial beliefs may then be updated in different ways depending on the experiences of principals in the different contractual settings, which is one of the aspects we want to study, as part of how initial trust and contract enforcement interact.

Subjects in the role of agents did not receive any example, nor were they informed that the subjects in the role of principals received such information. The instructions for agents did not differ in the high- and low-trust treatments, which rules out any direct impact on outcomes through an influence on agents. This illustrates the advantages of an experimental setting for varying only the principals' trust, defined as the principals' beliefs in the trustworthiness of the agents.

1.4. *Factorial Design*

We exposed all principals in a session to either the high-trust (HT) or the low-trust (LT) manipulation in the three contracting environments WEAK, MEDIUM and STRONG. We thus implemented a 3×2 factorial design, crossing contracting environment and trust manipulation. This is indicated in the labels that we use to refer to the treatments. WEAK-HT, for example, stands for the treatment in which we implemented the weak contracting environment and exposed the principals to the high-trustworthiness example. We conducted five markets for each of our six main treatments, as shown in Table 2.

1.5. *Additional Treatments*

We conducted the additional treatment STRONG-HT-Long, which is identical to STRONG-HT, except that the game lasted twenty-five periods rather than fifteen periods. The purpose of STRONG-HT-Long is to clarify the potential role of endgame effects.¹⁰

We also conducted the additional treatment TRANSPARENCY-LT, which is identical to STRONG-LT with one key difference. In STRONG-LT, it is common knowledge that principals can observe only the realised values of the agents with whom they interacted. In TRANSPARENCY-LT, principals can observe the realised values of all agents who have a

¹⁰ Endgame effects reliably occur in finitely repeated cooperation games such as the prisoner's dilemma and can typically be shifted into the future with a longer finite horizon (Embrey *et al.*, 2018). Similar effects can be observed in finitely repeated gift-exchange experiments (e.g., Brown *et al.*, 2004). Thus, we conjectured that we will also observe an endgame effect in treatment STRONG-HT. Since we are interested in the stable effects of trust in STRONG-HT (i.e., in periods in which the endgame effect is not operative), we wanted to examine whether we can extend the effect of exogenous increases in trust by simply increasing the number of periods.

contract in a given period. Principals record this information on a printed form at the end of each period, exactly as in STRONG-LT. The only difference is that instead of recording a single value (10 or 100), they must note up to seven values for the up to seven agents who have a contract in a given period. The treatment TRANSPARENCY-LT captures situations where contract offers can condition on even more information, for example on reference letters from previous employers. We used TRANSPARENCY-LT to explicitly demonstrate that sufficiently strong contract enforcement can substitute for high levels of trust.

We conducted two markets under the STRONG-HT treatment and five markets under the TRANSPARENCY-LT treatment, as shown in Table 2.

1.6. *Experimental Procedures*

We implemented a between-subject design, i.e., each subject participated in only one market in one treatment. Altogether we have thirty-seven markets, with seven principals and ten agents each. Hence, 629 subjects participated in our experiment. Subjects were mainly students from the University of Zurich and the Swiss Federal Institute of Technology in Zurich. Students majoring in economics or psychology were not eligible to participate.

All sessions were conducted in the computer laboratory of the Department of Economics at the University of Zurich in 2011 and 2012, except for the TRANSPARENCY-LT treatment, which took place in 2023. The study was computerised using the software z-Tree (Fischbacher, 2007), and participant recruitment, except for TRANSPARENCY-LT, was conducted using the software ORSEE (Greiner, 2015). Before the subjects entered the lab, they randomly drew a place card that specified at which computer terminal to sit. The terminal number determined a subject's role as either principal or agent, which remained fixed throughout the experiment.

Subjects received written instructions including comprehension questions, which had to be answered correctly before a session could begin.¹¹ A summary of the instructions was read aloud by the experimenter to generate common knowledge of the instructions. To familiarise participants with the market procedures, two practice periods were conducted before the main experiment. During these practice rounds, subjects only completed the first stage of the experiment, meaning that principals did not see payoffs and were unable to infer agents' actual effort decisions. Subjects did not earn any money during the practice periods.

Sessions lasted about 2.5 hours. Payoffs from the experiment, denominated in points, were converted into money at the rate of 10 points to CHF 1 at the end of a session. On average, subjects earned about CHF 48, which includes a show-up fee of CHF 20. The subjects received their payments privately.

1.7. *Manipulation Check*

Our experimental approach aims at inducing exogenous variation in the principals' beliefs about the agents' trustworthiness. Figure 2 provides a manipulation check for our six main treatments by showing the principals' expectations about the empirical relationship between offered wages and chosen effort. These expectations were elicited at the beginning of the experiment, after reading the instructions, but before entering the trading periods. We asked principals to predict

¹¹ We provide the experimental instructions in [Online Appendix C](#). Since the terms 'principal' and 'agent' are not in common usage among student subjects, the experiment was framed in terms of 'buyers' and 'sellers', and we spoke about 'price' and 'quality' instead of 'wage' and 'effort'.

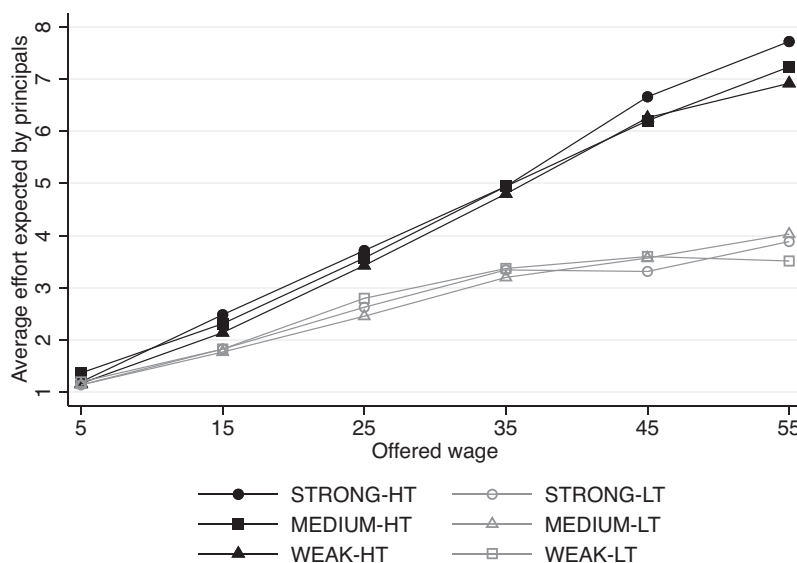


Fig. 2. *Manipulation Check.*

Notes: The black lines show the wage-effort relation that principals expect in the three high-trust treatments WEAK-HT, MEDIUM-HT and STRONG-HT. The grey lines show expectations in the three low-trust treatments WEAK-LT, MEDIUM-LT and STRONG-LT.

what they thought would be the average effort level chosen by agents, conditional on different possible offered wages.

The figure reveals exogenous belief variation in all contracting environments, WEAK-HT versus WEAK-LT, MEDIUM-HT versus MEDIUM-LT and STRONG-HT versus STRONG-LT. Principals expected significantly higher average effort levels across the range of wages when they had received the high-trustworthiness example rather than the low-trustworthiness example, and also expected significantly steeper relationships between wage and effort. Regressions confirm that the differences in average expected effort across the high- and low-trustworthiness examples were statistically significant at the 1% level, as were the differences in slopes, in all three treatment pairs (see [Online Appendix Tables A4–A5](#)). On the other hand, we cannot reject the hypotheses that the average expected effort, and the slopes of the wage-effort relations, are identical when comparing across treatments involving the low-trustworthiness example, and across treatments involving the high-trustworthiness example (see [Online Appendix Tables A6–A7](#)).

Since the agents do not receive historical examples about agents' trustworthiness in a previous experiment, random assignment should result in no differences in agents' beliefs across treatments involving high- versus low-trustworthiness examples. Indeed, we cannot reject the hypothesis that the agents' expectations about average effort, and the slope of the wage-effort relations, are identical within each of the treatment pairs (see [Online Appendix Tables A8–A9](#)). Since the agents indicate their 'homegrown' beliefs, we can compare these beliefs with the beliefs that the principals indicate. We find that the principals' beliefs in the high-trust treatments roughly correspond to the agents' homegrown beliefs. Principals who received the low-trustworthiness example thus have beliefs that are more pessimistic than homegrown beliefs.

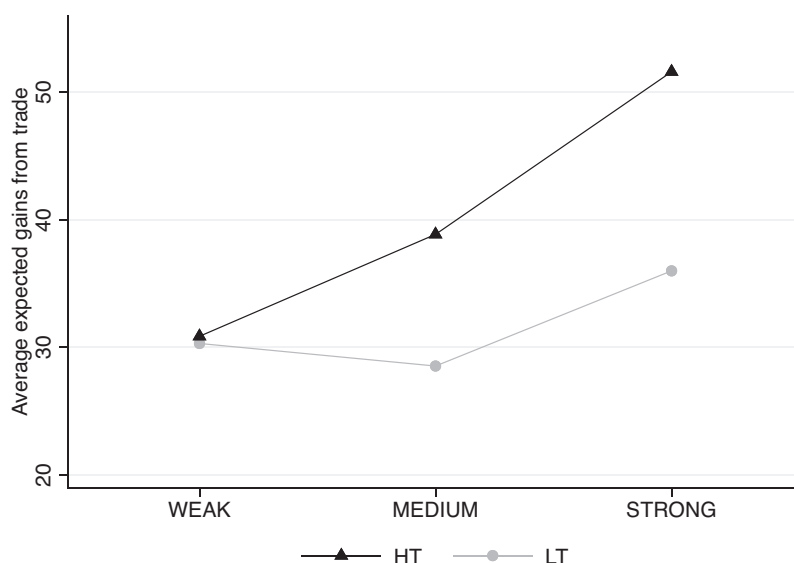


Fig. 3. *Expected Gains from Trade.*

Notes: The black line shows the average expected gains from trade in the high-trust environment in treatments WEAK, MEDIUM and STRONG. The grey line shows the respective average expected gains from trade in the low-trust environment.

2. The Complementarity of Trust and Contract Enforcement

In this section, we present our main result on the complementarity between trust and contract enforcement for the expected gains from trade. Figure 3 shows the average expected gains from trade in our three contracting environments, comparing the high-trust and low-trust environments.

First, Figure 3 reveals that an exogenous increase in trust has no significant effect on expected gains from trade in contracting environment WEAK, where principals are not obliged to pay the promised wage (Wilcoxon rank-sum test, $p < .92$). This is, in our view, an interesting finding because it suggests that, under weak legal and economic contract enforcement institutions, higher trust alone might not increase trading efficiency.

Second, the figure reveals that the effect of trust on expected gains from trade is positive under improved contract enforcement conditions. The causal effect of trust is positive and statistically significant in contracting environment MEDIUM, where the principals are contractually obliged to pay the offered wage, and in environment STRONG, where contract terms can be conditioned on signals of past agent performance (Wilcoxon rank-sum tests of market averages, $p < .02$, $p < .01$, respectively).¹²

Third, the figure shows that the effect of trust on expected gains from trade is largest in contracting environment STRONG.¹³ Corresponding regressions show that the differences in the impact of trust are significantly different across the institutional environments (for MEDIUM

¹² All Wilcoxon rank-sum tests reported in the paper are based on market averages, i.e., taking markets as the unit of independent observation, to allow for potential interdependence between observations from the same market.

¹³ Figure 3 obscures the fact that the effect of trust on gains from trade is transitory in MEDIUM, an aspect to which we will return later in the paper.

versus STRONG, this is true when excluding an endgame effect in STRONG, which we discuss in more detail later in the paper).¹⁴

Viewed differently, Figure 3 also illustrates that the effects of exogenous improvements in the contracting environment are trust dependent. While improvements in contract enforcement induce a large increase in expected gains from trade in the high-trust environment, much smaller effects, if any, are observed in the low-trust environment. This again indicates a complementarity between trust and the contract enforcement environment for trading efficiency. It is not sufficient to merely improve the contract enforcement environment or the actors' beliefs in the trustworthiness of their trading partners. The greatest increase in gains from trade emerges when both contract enforcement and trust are simultaneously enhanced. We summarise these observations in our first result.

RESULT 1. *Trust and contract enforcement can be complements with regard to gains from trade. While an exogenous increase in trust has no effect on expected gains from trade in contracting environment WEAK, it induces higher expected gains from trade in MEDIUM and STRONG, with the increase being largest in STRONG. Likewise, while an exogenous improvement in contract enforcement causes no or little increase in expected gains from trade in the low-trust environment, it induces substantially higher expected gains from trade in the high-trust environment.*

3. Mechanisms

The previous section established that trust and contract enforcement can be complementary for gains from trade between principals and agents. In this section, we study the mechanisms underlying this finding, by analysing the behaviours of principals and agents in more detail.

3.1. Weak Contract Enforcement

Principals are not obliged to pay the offered wages, and agents are not obliged to choose the requested effort levels in contracting environment WEAK. Recall that interactions are one-shot, IDs are randomised across periods, and principals and agents, after agreeing on a contract, simultaneously choose their actual wages and effort levels, respectively.

Figure 4(a) reveals that the principals, on average, do not honour their promises. The dashed lines show the average offered wages over the fifteen periods of the experiment in the high-trust treatment WEAK-HT (filled squares) and in the low-trust treatment WEAK-LT (open squares). The solid lines show the average actual wages in WEAK-HT (filled squares) and WEAK-LT (open squares). While offered wages even increase over the course of the fifteen periods, reaching values above 50 in WEAK-HT and above 40 in WEAK-LT, actual wages paid decrease to values below 10 in both treatments. Principals honoured the promised wages in only 13% and 11% of all transactions in WEAK-HT and WEAK-LT, respectively.

¹⁴ See the interaction terms HT×WEAK and HT×STRONG in column (5) in Table 3 where effort, being a sufficient statistic for expected gains from trade, is regressed on our contractual environments (WEAK, MEDIUM, STRONG), the trust level (HT, LT) and their interactions, with LT × MEDIUM as the omitted category. The positively significant coefficient of the interaction term HT×STRONG in regression (5) ($p < .04$), which excludes the final five periods of STRONG, indicates that the positive impact of HT on effort is larger in STRONG than MEDIUM. The negatively significant coefficient of HT×WEAK ($p = .011$) indicates that the impact of HT is smaller in WEAK than MEDIUM. Results are similar although less precise if we exclude the final five periods in all treatments, as shown in regression (6) of Table 3.

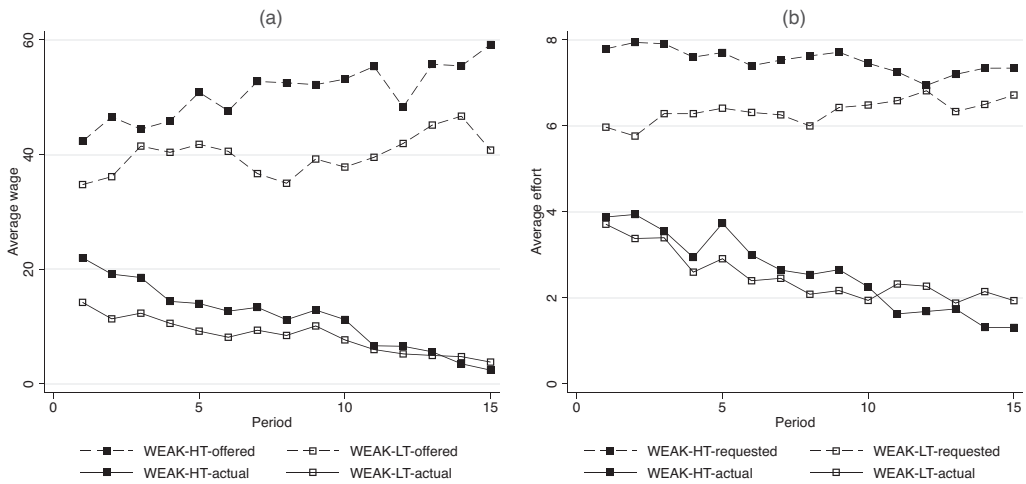


Fig. 4. *Promised and Actual Wages, and Requested and Actual Effort, in Contracting Environment WEAK.* Notes: In (a) the filled squares show offered (dashed line) and actual (solid line) wages in the high-trust environment. The open squares show offered (dashed line) and actual (solid line) wages in the low-trust environment. In (b) the filled squares show requested (dashed line) and actual (solid line) effort levels in the high-trust environment. The open squares show requested (dashed line) and actual (solid line) effort levels in the low-trust environment.

Actual wages are, on average, 11.6 in WEAK-HT and 8.5 in WEAK-LT (Wilcoxon rank-sum test, $p = .095$). While there is a small and marginally significant effect of an exogenous increase in trust on the average actual wages paid, Figure 4(a) reveals that this effect is driven by the earlier periods only. Regression (1) in [Online Appendix Table A1](#) confirms that the difference in actual wages paid between WEAK-HT and WEAK-LT diminishes over time ($p < .1$). We summarise these observations in our next result.

RESULT 2. *In contracting environment WEAK, the principals promise to pay high wages to the agents, but rarely honour their promises, both in the high-trust and in the low-trust environments. Actual wages are low in both environments. While promised wages are always higher in the high-trust environment than in the low-trust environment, actual wages converge towards the same level in both environments.*

Result 2 shows that the agents have little reason to believe that the principals pay the promised high wages. As a consequence, the agents—regardless of whether they are reciprocal or selfish—have no reason to provide the effort levels requested by the principals in the contracts. Indeed, Figure 4(b) shows that the agents in contracting environment WEAK do not deliver the effort levels that the principals requested. The two dashed lines show the average effort levels requested by the principals over the fifteen periods in WEAK-HT (filled squares) and WEAK-LT (open squares). The two solid lines show the average actual effort levels delivered by the agents in WEAK-HT (filled squares) and WEAK-LT (open squares). While requested effort levels range between 7 and 8 in the high-trust environment and between 6 and 7 in the low-trust environment, delivered effort levels decrease to values around 2 in both environments. Agents delivered the

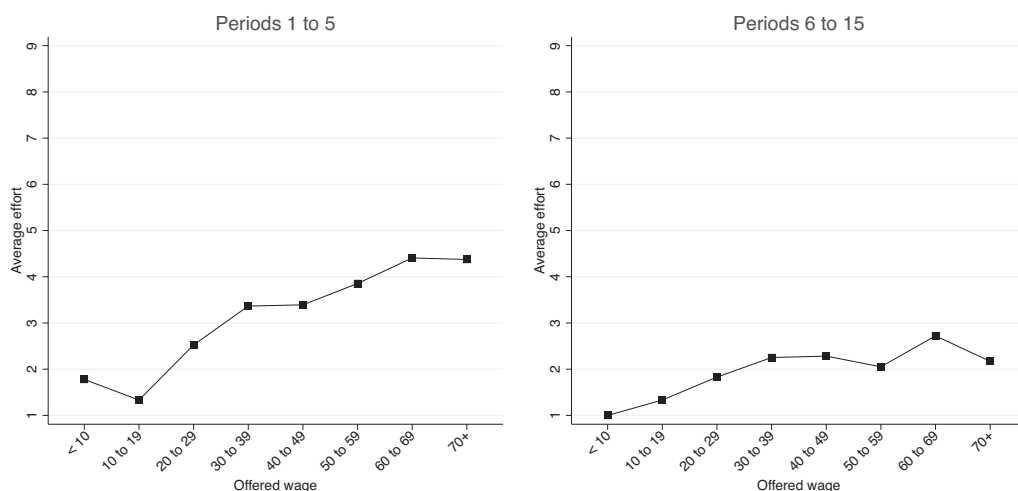


Fig. 5. *Promised Wages and Actual Effort Levels in Contracting Environment WEAK.*

Notes: The left panel shows the relation between promised wages and actual effort levels in periods 1 to 5; the right panel shows the relations for periods 6 to 15. Data from WEAK-HT and WEAK-LT pooled.

requested effort level only in 7% and 16% of all transactions in WEAK-HT and WEAK-LT, respectively.¹⁵

Figure 4(b) shows that agents' effort is essentially identical in WEAK-HT and WEAK-LT. Effort levels in WEAK-HT tend to start at slightly higher levels than in WEAK-LT, but also decline slightly more over time (see the small, but significantly negative coefficient on $HT \times Period$ in regression (1) in [Online Appendix Table A2](#); $p < .05$). No difference exists in average actual effort levels in WEAK-HT and WEAK-LT (2.6 and 2.5; Wilcoxon rank-sum test, $p = .75$). We summarise this finding next.

RESULT 3. *In contracting environment WEAK, the agents rarely deliver the requested effort level, both in the high-trust and in the low-trust environments. Actual effort levels are low and very similar in both trust environments.*

A key feature of contracting environment WEAK is that the principals cannot commit to paying high wages and, therefore, the agents may not consider high offered wages as a credible promise. Thus, for agents with a reciprocity motive, the desire to reciprocate to high wage offers with high effort choices is undermined. However, the agents may initially, i.e., during the early periods, not know the extent to which the principals' wage offers are credible. In fact, a closer look at the data reveals that traces of reciprocity exist, even in contracting environment WEAK. This holds, in particular, in early periods of the experiment, when reciprocal agents might not have fully realised that promised wages are rarely paid by the principals. Figure 5 shows actual effort levels as a function of *promised* wages. The left panel shows the relation for periods 1 to 5, while the right panel shows the relation for periods 6 to 15. The figure reveals that agents, on

¹⁵ Of agents choosing less than the requested effort in the WEAK conditions, roughly 60% choose the minimum effort, but about 40% choose an effort level between the minimum and requested effort. This is consistent with agents not expecting to receive the offered wage, but some reciprocal agents choosing non-minimal effort levels to reciprocate an anticipated actual wage that is below the offered wage, but not the minimum possible wage.

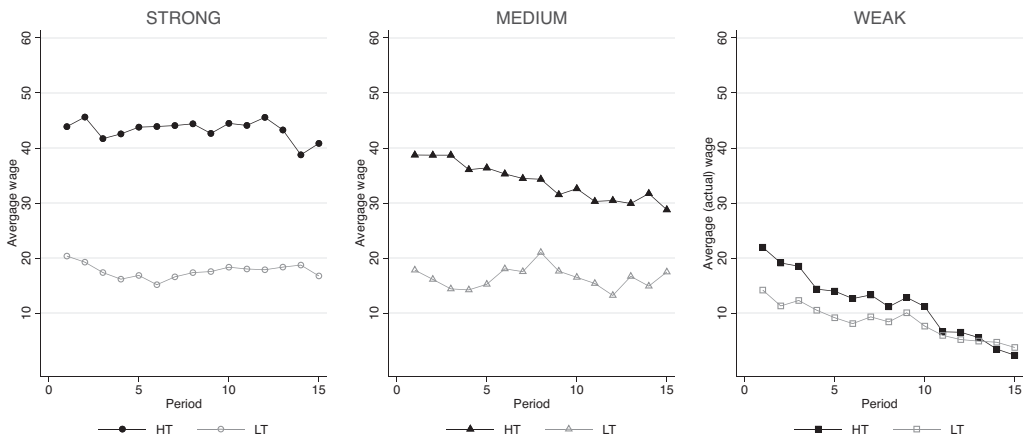


Fig. 6. Actual Wage Levels over Time.

Notes: The black lines show average wages in sessions with the high-trust treatments and the grey lines show average wages in sessions with the low-trust treatments.

average, responded to high wage offers with somewhat higher effort levels in early periods, but that the relation is substantially flatter in later periods (the Spearman correlation is 0.27 in periods 1 to 5, compared to 0.08 in periods 6 to 15).¹⁶ Thus, the lack of a legal commitment opportunity for the principals in contracting environment WEAK together with the agents' experience during the early periods that the principals rarely honour their promises appears to have weakened the reciprocal agents' responses to high promised wages, explaining the dynamic of falling actual effort levels over time.

3.2. Medium Contract Enforcement

Contract enforcement is strengthened in contracting environment MEDIUM relative to contracting environment WEAK because the principals are contractually obliged to pay the offered wages in contracting environment MEDIUM. Agents remain free to choose whatever effort they like. The credible commitments to higher wages, however, may induce reciprocal agents to provide higher effort levels when wages are higher.

We first examine the wage behaviour of the principals. The middle panel of Figure 6 shows actual wages in contracting environment MEDIUM. In contrast to contracting environment WEAK (shown again in the right panel for ease of comparison across treatments), the exogenous increase in trust is associated with a significant, though unstable, difference in actual wages. Wages are 33.9 on average in MEDIUM-HT and 16.5 in MEDIUM-LT. This difference of about 17 points is statistically significant (Wilcoxon rank sum test, $p < .01$). Regression analysis confirms that the impact of the exogenous increase in trust is significantly larger in MEDIUM than in WEAK; see the large negative coefficient on the interaction term $HT \times WEAK$ in regression (1) in Table 3. However, the effect of the exogenous trust increase in MEDIUM becomes smaller

¹⁶ The difference in slopes is also statistically significant. This can be seen in a panel regression of effort on offered wage, a dummy variable for period > 5 , and an interaction term between this dummy and offered wage (with random effects for principals, and bootstrapped SEs clustering on session); the interaction term is highly significant and negative ($p < .01$).

Table 3. *The Impact of Trust and Contractual Environments on Actual Wages and Effort Levels.*

	Actual wage			Effort		
	All periods (1)	No endgame (2)	Periods < 11 (3)	All periods (4)	No endgame (5)	Periods < 11 (6)
HT	17.40*** [3.03]	17.40*** [3.03]	18.75*** [2.38]	1.40*** [0.44]	1.40*** [0.44]	1.46*** [0.48]
WEAK	−8.06*** [1.22]	−8.06*** [1.22]	−6.81*** [1.60]	0.23 [0.24]	0.23 [0.24]	0.33 [0.28]
STRONG	1.10 [2.65]	0.94 [2.67]	0.48 [2.74]	1.03*** [0.32]	1.13*** [0.25]	1.03*** [0.29]
HT x WEAK	−14.24*** [3.35]	−14.24*** [3.35]	−14.02*** [3.12]	−1.32** [0.52]	−1.32** [0.52]	−1.06* [0.59]
HT x STRONG	8.33* [5.02]	8.89* [4.91]	7.54* [4.55]	0.75 [0.58]	1.11** [0.55]	1.05* [0.59]
Constant	16.49*** [0.95]	16.49*** [0.95]	16.95*** [1.08]	2.28*** [0.19]	2.28*** [0.19]	2.38*** [0.24]
Observations	3,119	2,771	2,083	3,119	2,771	2,083

Notes: Panel regression estimates with random effects for principals. Bootstrapped SEs clustered on markets (thirty clusters) are reported in brackets. Columns (1) to (3) present regressions explaining actual wages as a function of contracting and trust environments. In columns (4) to (6) the dependent variable is actual effort levels. The estimations use data from all three contracting environments with LT × MEDIUM as the omitted category. HT is a dummy variable indicating the respective high-trust environment. Thus, HT measures the impact of the high-trust environment under medium contract enforcement. WEAK measures the impact of weak contract enforcement relative to MEDIUM in the low-trust environment and STRONG measures the effect of strong contract enforcement relative to MEDIUM in the low-trust environment. HT × WEAK and HT × STRONG give the differential effect of high trust in these respective contracting environments relative to the impact of high trust in MEDIUM. Columns (2) and (5) only use periods 1 to 10 for contracting environment STRONG to eliminate the endgame effect that is present in that environment. Columns (3) and (6) check robustness to only using periods 1 to 10 for all treatments. ***, **, * Significance at the 1%, 5% and 10% levels, respectively.

over time, showing that the initial impact of trust on actual wages is steadily declining (see the interaction term HT × Period in regression (2) in [Online Appendix Table A1](#); $p < .03$). We summarise this observation in our next result.

RESULT 4. *In contracting environment MEDIUM, an exogenous increase in trust induces an initial increase in principals' wage payments, but the wage difference across trust environments declines over time.*

We now turn to the behaviour of agents. Figure 7 shows effort levels over time in our main treatments. The middle panel shows that average effort is significantly higher in MEDIUM-HT compared to MEDIUM-LT (3.7 and 2.3, respectively; Wilcoxon rank-sum test, $p = .02$). Regression analysis confirms that the difference in effort levels between the high-trust and the low-trust environments is significantly larger in contracting environment MEDIUM than in contracting environment WEAK; see the negative coefficient of HT × WEAK in regression (4) of Table 3. However, the effect of the exogenous increase in trust is declining over time. This time trend is significant, as indicated by regression analysis, and is driven by the decline in average effort levels in MEDIUM-HT (see the interaction term HT × Period in regression (2) in [Online Appendix Table A2](#); $p < .01$).¹⁷ We summarise this finding in the following result.

¹⁷ The small remaining difference in effort between HT and LT in MEDIUM is no longer significant in the final period (Wilcoxon rank-sum test, $p = .34$), corroborating that the effect of the exogenous trust increase is merely transitory in the contracting environment MEDIUM.

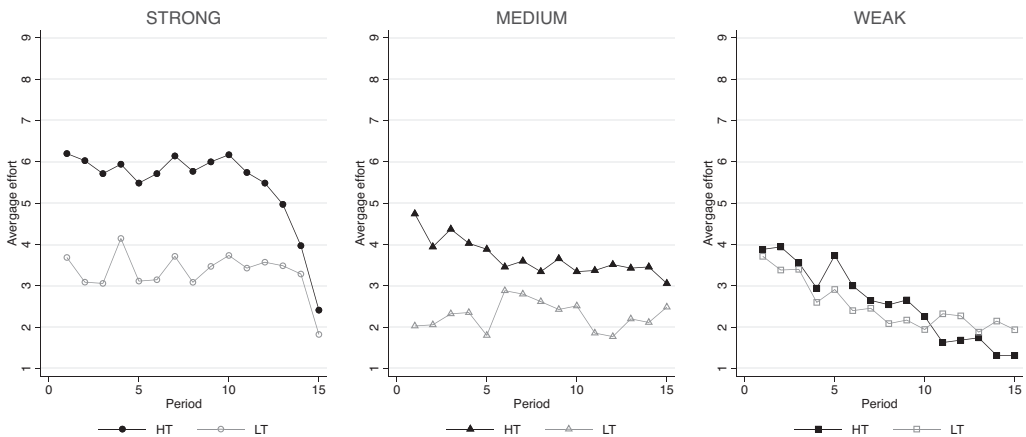


Fig. 7. *Actual Effort Levels over Time.*

Notes: The black lines show average effort in sessions with the high-trust treatments and the grey lines show average effort in sessions with the low-trust treatments.

RESULT 5. *In contracting environment MEDIUM, the higher wage payments in the high-trust environment are associated with higher average effort levels. However, the difference in effort levels between the high-trust and the low-trust environments declines over time.*

We hypothesised that the lack of wage commitment among principals in contracting environment WEAK undermined the possibility to elicit high effort levels from reciprocal agents by promising high wages. In contracting environment MEDIUM, by contrast, wage promises are credible. Therefore, we hypothesise that the higher effort levels observed in MEDIUM-HT may reflect the ability of principals to elicit reciprocal responses from some agents, although the randomly shuffled IDs prevent strategies of screening for reciprocal agents.¹⁸ Indeed, Figure 8(a) shows that the effort levels delivered by agents in MEDIUM-HT are, on average, responsive to the wages paid by principals. This is unlike in WEAK-HT, where the effort-wage relation is essentially flat. Corresponding regressions relating effort to actual wages confirm that this difference in slopes is statistically significant (see [Online Appendix Table A3](#)). In regression (1), the coefficient on actual wages is significantly positive ($p < .01$), while the interaction term between actual wages and WEAK-HT is significantly negative and of a similar absolute size ($p < .01$), indicating that the effort-wage relation is flat in WEAK-HT. The positive average effort response to wages in MEDIUM-HT is consistent with some agents being committed by their intrinsic preferences for reciprocal behaviour. However, since principals reduce their wage offers over time in MEDIUM-HT, the reciprocally motivated agents respond by reducing their effort levels over time. The next result summarises these findings.

¹⁸ Screening requires endogenously choosing to interact with the same agent repeatedly. The frequency of interactions in which principals and agents continue to interact is very low in treatments with randomly re-shuffled IDs; about 9% of interactions involve the principal and agent having interacted in the previous period. Considering the most permissive definition of repeat interactions, based on whether the same principal and agent have ever interacted, we find that 47% of interactions involve the same principal and agent having met at any time previously. By contrast, in treatments with fixed IDs, about 33% of interactions are a continuation of an interaction in the previous period, and about 64% of interactions are a repeat of an interaction at some point in the past.

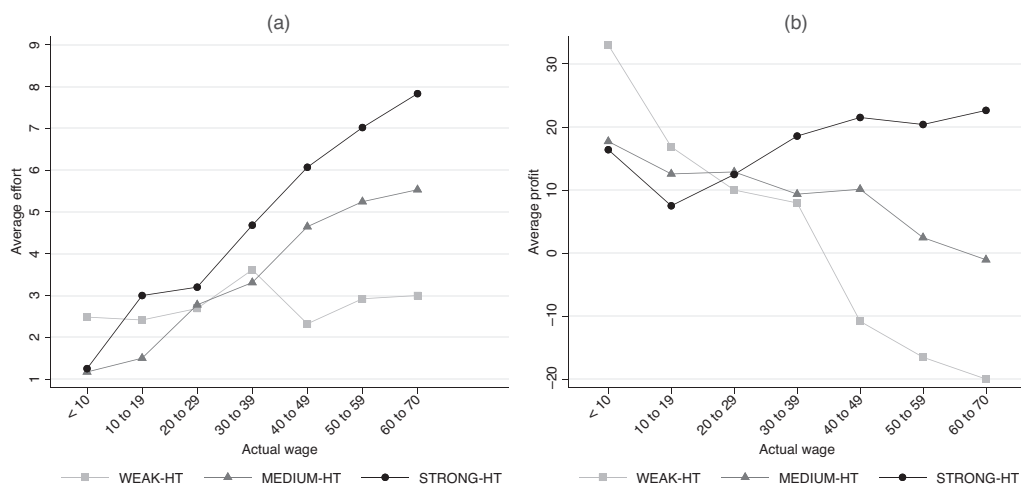


Fig. 8. *Effort-Wage and Profit-Wage Relations in High-Trust Environments.*

Notes: Panels (a) and (b) respectively show the average effort levels and average profits for the principal, for different actual wage levels in the high-trust conditions of contracting environments WEAK, MEDIUM and STRONG. Promised and actual wage payments coincide in MEDIUM and STRONG. The figure excludes the final five periods in contracting environment STRONG to show the relations absent the endgame effect. The figure also excludes categories of actual wages above 70 as there are too few observations for meaningful comparisons (e.g., only one observation in this range for WEAK).

RESULT 6. *Unlike in contracting environment WEAK-HT, in contracting environment MEDIUM-HT, agents respond to higher actual wages with higher effort levels on average. However, as principals reduce wages over time in MEDIUM-HT, effort levels also fall.*

The tendency for principals to reduce wages over time in MEDIUM-HT is understandable because the average effort response to wages (Figure 8(a)) is not strong enough to make high wages profitable for principals. As shown in Figure 8(b), average profits for principals in MEDIUM-HT decline with actual wages. This decline is steep for wages above 50, but relatively flat for wages in the range 10 to 39, which may explain why principals only slowly learned that lowering wages in this range is more profitable. Corresponding regressions show that the relationship of profits to wages is negative in MEDIUM-HT (see the coefficient on Actual wage in regression (3) in [Online Appendix Table A3](#), $p < .01$). As expected, panel (b) shows that paying non-minimal wages strongly reduces profits in WEAK-HT because there is no reciprocal response of effort to actual wages in this environment. Corresponding regressions show that the difference in slopes for WEAK versus MEDIUM is statistically significant (see the interaction term Actual wage $\times$ WEAK-HT in regression (3) in [Online Appendix Table A3](#), $p < .01$).

A reason for the modest average response of effort to wages in MEDIUM-HT is heterogeneity in agent behaviour. Some agents exhibit a strong reciprocal response to high wages, while others behave selfishly, choosing low effort levels even when wages are high. Indeed, as shown in the upper panels of Figure 9, for high (50+) and medium (25 to 50) wages, there is substantial heterogeneity in agent effort choices in MEDIUM-HT, with effort choices spanning the entire range. The effort distributions have modes at minimum effort and at relatively high effort (about

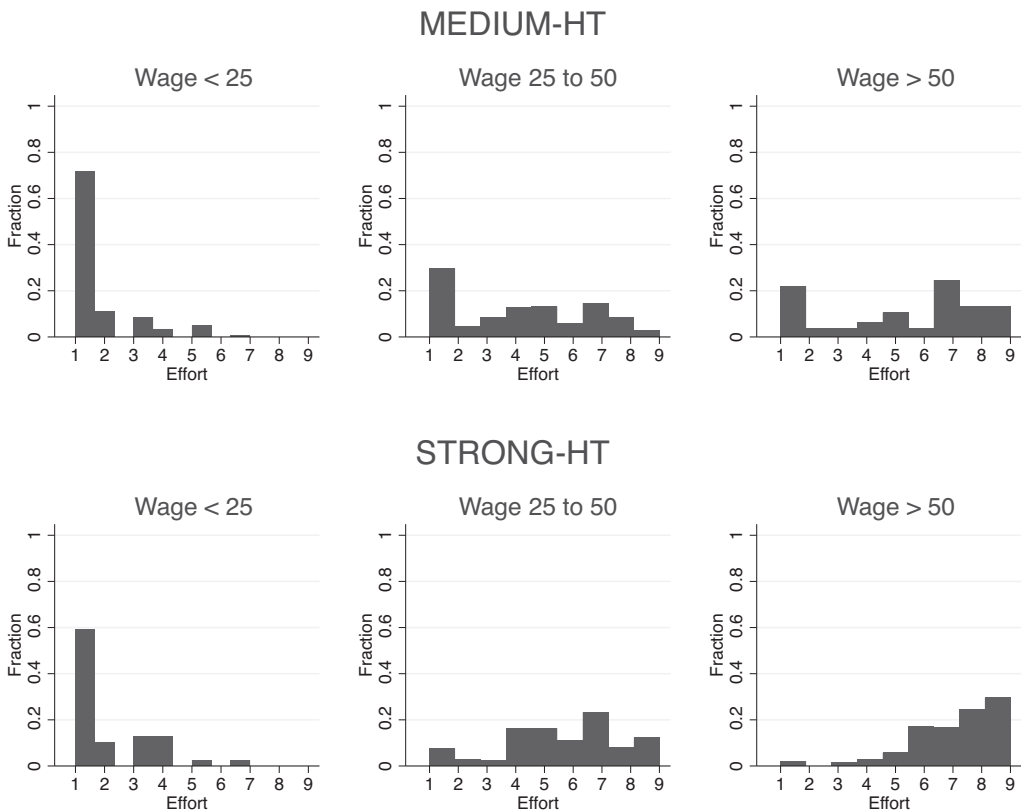


Fig. 9. Distributions of Agents' Effort Choices Conditional on Wage Ranges in the High-Trust Environments.

7), respectively.¹⁹ This heterogeneity means that paying a high or even medium wage runs the risk of getting an agent who chooses low or minimum effort, thereby generating a large loss. As shown in Figure 8(b), principals are better off paying relatively low wages, even though this results in low effort levels from both reciprocal and selfish agents (see the upper left-hand panel of Figure 9). These findings are summarised in the next result.

RESULT 7. *Paying high wages is not profitable in contracting environment MEDIUM-HT because the average effort response is too weak. The modest average response reflects underlying heterogeneity among agents, with some agents responding strongly, but many exhibiting a weak or zero response.*

¹⁹ Another indication of heterogeneity is provided by [Online Appendix Figure A1](#), which shows that principals in different market sessions of MEDIUM-HT experienced slightly steeper or flatter responses of effort to actual wage, consistent with some heterogeneity in the number and strength of reciprocal agent types across sessions. There is a clear tendency, however, for these relationships to be flatter in MEDIUM-HT than those observed in sessions of STRONG-HT, and steeper than those observed in sessions of WEAK-HT.

3.3. *Strong Contract Enforcement*

3.3.1. *Aggregate wages and effort levels*

Contract enforcement opportunities are further enhanced in contracting environment STRONG. As in contracting environment MEDIUM, principals are obliged to pay the offered wages. However, enforcement is expanded further in STRONG because fixed IDs allow principals to condition their contract offers on signals about their agents' past effort choices.

The left panel of Figure 6 shows actual wages in STRONG. Unlike in MEDIUM, the exogenous increase in trust is associated, not only with a significant, but also with a stable increase in actual wages. Average wages are 43.3 in STRONG-HT and 17.7 in STRONG-LT, a difference that is highly significant (Wilcoxon rank sum test, $p < .01$). Regression analysis confirms that the treatment difference is stable over time, with no statistically significant time trend for the treatment difference in wages (see $HT \times \text{Period}$ in regressions (3) and (4) in [Online Appendix Table A1](#)).

Notably, wages are similar in the low-trust environments in STRONG and MEDIUM, whereas this is not the case in the high-trust environments. The average wages in STRONG-LT and MEDIUM-LT are not significantly different (17.7 versus 16.5; Wilcoxon rank sum test, $p = .84$) and remain similar throughout the fifteen periods of the experiment. In contrast, while wages are initially only slightly lower in MEDIUM-HT than in STRONG-HT, the gap between these treatments strongly increases over time as wages in MEDIUM-HT steadily decline. Overall, average wages are marginally significantly lower in MEDIUM-HT than STRONG-HT (33.9 versus 43.3, respectively; Wilcoxon rank sum test, $p = .095$). Regression analysis shows that the overall impact of the exogenous trust increase on wages is marginally significantly larger in STRONG than in MEDIUM (see the interaction term $HT \times \text{STRONG}$ in regressions (1) to (3) of Table 3, $p < .1$).

Turning to agent behaviour, the left panel of Figure 7 shows the average effort levels over time in contracting environment STRONG. The figure reveals that effort is substantially higher in STRONG-HT than in STRONG-LT. Average effort is 5.5 in STRONG-HT and 3.3 in STRONG-LT, a difference that is significant (Wilcoxon rank sum test, $p = .016$) and stable throughout the experiment, except for an endgame effect in the final periods.

To determine whether the decline in effort towards the end of the game (i.e., in periods 11–15) is indeed an effect tied to the end of the game, as opposed to a time trend in the effect of high trust that happens to start after ten periods, we conducted two markets of a control treatment, labelled STRONG-HT-Long. Treatment STRONG-HT-Long is identical to treatment STRONG-HT, except that it lasted twenty-five periods rather than fifteen periods. Figure 10 shows that effort remains high for ten additional periods in STRONG-HT-Long, declining only as the longer game approaches its end. Extending the game thus prolongs the length of the stable effect of high trust, moving the decline in effort to the end of the longer game. The endgame effect is thus relatively unimportant economically because it influences behaviour only in the final periods of play, and its occurrence can be shifted into the future by extending the duration of the game.

Note that treatment STRONG-HT-Long replicates the high level of gains from trade observed in STRONG-HT. Average effort in STRONG-HT-Long is 5.6, almost identical to (and not significantly different from) the average effort of 5.5 in STRONG-HT. Regression analysis confirms that, excluding the endgame effect for effort, there is no statistically significant time trend in STRONG-HT for the high- versus low-trust treatment difference in effort levels (see the near-zero coefficient of $HT \times \text{Period}$ in regression (4) in [Online Appendix Table A2](#)). We summarise these effects of exogenous trust on wages and effort in the following result.

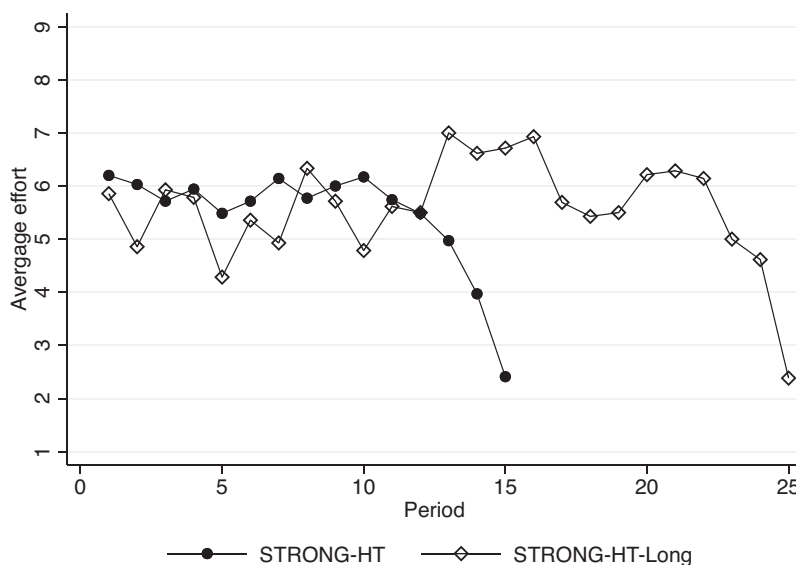


Fig. 10. *Robustness Check Verifying the Endgame Effect in the Final Periods of STRONG-HT.*
Notes: Average effort levels are more volatile over the course of the experiment in STRONG-HT-Long than in STRONG-HT because we conducted only two markets in STRONG-HT-Long, not five as in STRONG-HT.

RESULT 8. (a) *In contracting environment STRONG, principals pay substantially higher wages in the high-trust environment compared to the low-trust environment. This wage difference is stable over time and, after the initial periods, significantly larger than the wage difference in contracting environment MEDIUM. Wage levels are similar over time in STRONG and MEDIUM when trust is low.*

(b) *The large positive effect of the exogenous trust increase on wages is associated with a large increase in effort. Principals in STRONG-HT pay high wages and receive high effort on average. Conversely, principals in STRONG-LT pay low wages and receive low effort on average.*

Comparing the effect of the exogenous trust increase in contracting environments STRONG and MEDIUM, we find that the difference in average effort between STRONG-HT and STRONG-LT (2.1) is larger than the difference between MEDIUM-HT and MEDIUM-LT (1.4). Regression analysis confirms that the effect of trust is significantly larger in STRONG than in MEDIUM; see the interaction term $HT \times Strong$ in regression (5) of Table 3, which excludes the endgame effect in STRONG-HT.²⁰ This stronger impact of trust on effort levels explains the larger impact of trust on gains from trade in STRONG compared to MEDIUM.

3.3.2. Screening strategies

These findings raise the question: what is the mechanism underlying the higher wages and effort levels in STRONG-HT compared to MEDIUM-HT? Additionally, why do principals in

²⁰ Regression (6) in Table 3 checks robustness to also excluding the final five periods of MEDIUM-HT; we see a very similar point estimate for $HT \times STRONG$ to that observed in regression (5), and the coefficient is still marginally significant; the larger SEs are understandable due to the reduction in sample size.

STRONG-LT behave similarly to those in MEDIUM-LT? We hypothesise that principals in STRONG-HT are optimistic about the prevalence of agents who reciprocate higher wages with higher effort and utilise the richer contracting environment to screen for such agents. For example, principals might pay medium wages in initial interactions with agents, and based on a positive signal, continue the relationship with similar or even higher wages. If this screening is successful, principals in STRONG-HT can selectively pay high wages only to agents who respond strongly, potentially making high wages profitable, unlike in MEDIUM-HT. In STRONG-LT, principals may be too pessimistic about the prevalence of agents who respond to high wages with high effort. Consequently, they avoid screening and instead pay low wages and elicit low effort levels, similar to the behaviour observed in MEDIUM-LT.

A precondition for screening to be a meaningful strategy is heterogeneity in how agents respond to a given wage. We saw in MEDIUM-HT how agents respond when wages are exogenous to agent type by construction: for low wages, almost all agents respond with low effort, but for high and medium wages, there is substantial heterogeneity (top panels of Figure 9). This heterogeneity implies that there may be value to principals in STRONG-HT in screening and assigning wages endogenously based on signals about agent behaviour. Furthermore, the way to identify agents who will reciprocate is to pay medium or high wages, as almost all agents choose low effort at low wages. Notably, screening with high wages may be riskier than using medium wages.

Looking at the wage offers of principals in STRONG-HT in initial and in later interactions with agents, we find behaviour consistent with such screening strategies. Principals in STRONG-HT tend to choose medium wages right from the outset in initial interactions with agents, on average paying a wage of 41. If a principal seeks out the same agent again—a sign of a successful previous interaction—the average wage tends to be even higher, 50 on average, thus moving into the range of high wages.²¹ Regressions confirm that in STRONG-HT, wages are significantly higher in later (private offer) interactions with an agent than in the initial interaction ($p < .01$).²² This pattern is consistent with principals moving to high wages once they have seen positive signals. In STRONG-LT, by contrast, principals start with low wages in initial interactions, 13 on average. If principals in STRONG-LT interact with the same agent again, the average wage is higher, at 27, but still not in the high range. The reluctance of principals in STRONG-LT to attempt screening is understandable if they hold a pessimistic view regarding the prevalence of agents who exhibit reciprocal behaviour.

The tendency for principals in STRONG-HT to pay medium wages in initial interactions, when they do not know anything about an agent, and high wages only after positive signals, is also apparent in the distributions of effort choices conditional on wage ranges. As shown in the middle bottom panel of Figure 9, there is substantial heterogeneity in agent effort choice when principals choose to pay medium wages in STRONG-HT (the figure excludes the endgame periods, so this heterogeneity is not driven by endgame effects). Indeed, principals face a similar heterogeneity as observed for such wages in MEDIUM-HT. For high wages, however, there is a very different pattern in STRONG compared to MEDIUM (compare the top and bottom right-hand panels of Figure 9). In STRONG-HT, when principals choose to pay high wages,

²¹ Another sign of screening strategies in STRONG-HT is shown in [Online Appendix Figure A2](#). High (above-median) first-period effort is associated with higher future average wages in STRONG-HT, but does not lead to higher future wages in MEDIUM-HT or WEAK-HT. This provides further evidence that principals in STRONG-HT are screening agents based on their past effort levels, and giving high wage offers to those who performed well.

²² We regress wages on a dummy variable for later interactions using a private offer, with the omitted category being initial interactions (with a private or public offer). The dummy variable is statistically significant. The estimation method is a panel regression with random effects on principals, and bootstrapped SEs clustering on session.

effort levels are almost uniformly high, with 88% of effort levels being 6 or higher, and the unique mode at a maximum effort level of 9. This indicates that principals are successful in targeting their high wage offers to reciprocal agents, which contrasts with MEDIUM-HT, where principals who pay high wages without the possibility of screening face the full range of effort outcomes and frequently experience relatively low effort levels. Regressions confirm that effort levels are significantly higher for high wages in STRONG-HT versus MEDIUM-HT ($p < .01$).²³ Turning to low wages in STRONG-HT, we see that effort levels are almost always low (left-hand bottom panel of Figure 9). This is the main type of wage paid by principals in STRONG-LT, and in STRONG-LT the effort response is similarly low.²⁴ This helps explain why principals in STRONG-LT experience low effort levels and why outcomes are similar to MEDIUM-LT, where principals also mainly pay low wages.

There is also direct evidence that principals in STRONG-HT condition contract terms on past signals, consistent with screening for agents who choose high effort levels in response to high wages. Specifically, if we regress the probability that an agent is ‘fired’ (i.e., was hired last period, but does not receive a private offer in the current period) on the previous period’s effort signal (controlling for previous wage), we find that in STRONG-HT a negative effort signal increases the probability of ‘firing’ by 28 percentage points ($p < .01$).²⁵ Similarly, among those who are rehired deliberately (using a private offer) despite a negative effort signal (i.e., realisation of the low value level) in $t - 1$, the probability of a lower wage offer increases by 16 percentage points ($p < .01$).²⁶

A sign that these strategies allow principals in STRONG-HT to be successful in screening would be a stronger relationship of effort to wage offers in STRONG-HT compared to MEDIUM-HT. As seen in Figure 8(a), the effort-wage relation is indeed steeper in STRONG-HT than in MEDIUM-HT, and corresponding regressions confirm that the difference is statistically significant (see the interaction term Actual wage $\times$ STRONG-HT in regressions (1) and (2) in [Online Appendix Table A3](#); $p < .05$). This again suggests that, when principals pay high wages in STRONG-HT—typically in later interactions with an agent after receiving positive signals—they are successfully targeting agents who respond with high effort levels and avoiding those who choose low effort levels.²⁷

²³ Results are based on regressing effort on a treatment dummy for STRONG-HT, conditioning on wage greater than 50 (random effects on principal with bootstrapped SEs clustering on market session). The coefficient on the treatment dummy is large and significant at the 1% level whether or not we include periods 11 to 15 for STRONG-HT.

²⁴ For low wages, average effort is approximately 2 in both STRONG-HT and STRONG-LT, and there is no statistically significant difference (regression of effort on a treatment dummy for HT, conditional on contract enforcement being strong and wage less than 25; random effects on principals with bootstrapped SEs clustering on market).

²⁵ About 50% of offers are private in STRONG-HT. Of these, about 57% are first-time private offers, but the remaining involve re-hiring an agent who was hired in the previous period with a private offer. This shows the screening process at work. In MEDIUM-HT, we also see roughly 40% private offers, but only 7% of these end up directed to an agent who was hired in the previous period with a private offer (and of these, most are likely accidental). This is as expected, given that random shuffling of IDs in MEDIUM-HT made screening essentially impossible. The fact that principals use private offers at all in MEDIUM-HT can be understood as an attempt to elicit a more reciprocal response from agents, with agents potentially appreciating being made an offer that is only available to themselves. Suggestive of this mechanism, in a regression of effort on price and an indicator for a private offer, the coefficient on private is positive and sizeable, indicating about half a point higher effort, albeit not statistically significant ($p < .25$).

²⁶ These results are based on a linear probability model with random effects for the principals, clustering on sessions and bootstrapped SEs. If we use, instead, a probit specification, we find very similar effects.

²⁷ Note that the endgame effect in STRONG-HT suggests that some agents who act reciprocal in pre-final periods start acting selfish at the end of the game. Thus, screening in pre-final periods may not be fully distinguishing agents with intrinsic reciprocity preferences from selfish agents who strategically imitate these types. Regardless, it can still be beneficial for principals to identify agents who act reciprocal in pre-final periods, as these are preferable to those who act selfish in pre-final periods.

The stronger relationship between effort and wages in STRONG-HT implies that paying high wages could be profitable. As seen in Figure 8(b), profits are indeed increasing with wages in STRONG-HT, while they decrease with wages in MEDIUM-HT. Corresponding regressions confirm that the profit response to wage increases is significantly better in STRONG-HT compared to MEDIUM-HT (see the interaction term Actual wage $\times$ STRONG-HT in regressions (3) and (4) in [Online Appendix Table A3](#); $p < .05$). Thus, unlike in MEDIUM-HT, principals in STRONG-HT can benefit from paying high wages, which generates high effort levels and high gains from trade. We summarise these findings in our next result.

RESULT 9. (a) *In contracting environment STRONG, high initial trust leads principals to engage in screening strategies to identify agents who respond strongly to high wages. These screening strategies are successful, resulting in a stronger relationship between effort and wages compared to contracting environment MEDIUM, making paying high wages more profitable.*

(b) *With low initial trust, principals in contracting environment STRONG do not try to screen for agents who respond to high wages. Instead, they pay low wages and elicit low effort levels similar to principals in contracting environment MEDIUM.*

4. Theory

In this section, we summarise our theoretical analysis of the principal-agent market game. We aim at providing an explanation for our main experimental finding, i.e., for why exogenous increases in trust have no effect on the gains from trade in contracting environment WEAK, only a temporary effect in contracting environment MEDIUM and a stable effect in contracting environment STRONG. The formal analysis can be found in [Online Appendix B](#).

To keep the analysis tractable, the game that we solve is a simplified version of the game used in the experiment. The model was constructed post hoc, and our paper does not aim to test the model. The goal of the theoretical analysis is simply to help organise and interpret the findings.

We model contracting environment WEAK as a one-shot simultaneous-move game between one principal, who chooses a wage, and one agent, who chooses an effort level. The game is further simplified by omitting the contract offer stage, which is pure cheap talk, and by assuming that all actions are binary. The principal is assumed to be profit maximising, while the agent can be either a selfish or a reciprocal type, but this type is unobservable to the principal.²⁸ Only reciprocal types would find it optimal to respond to a high wage with high effort in a one-shot interaction. However, since moves are simultaneous in contracting environment WEAK, there is a unique equilibrium in which wage and effort are always low. Importantly, this unique outcome is independent of the principal's belief about the share of reciprocal types in the agent population. Hence, even when our historical examples about the differential trustworthiness of agents in LT and HT changed a principal's belief about the share of reciprocal types, we predict that this has no effect on play, in line with our experimental finding that there is no difference between WEAK-LT and WEAK-HT.

We model contracting environment MEDIUM as a one-shot sequential-move game, where the principal first chooses a binding wage and the agent responds by choosing an effort after observing that wage. A reciprocal agent then indeed responds to a high wage with high effort,

²⁸ One could assume that there is also a positive fraction of reciprocal principals, but because preferences for reciprocal behaviour matter mainly for second movers (i.e., agents), we assume selfish principals. Reciprocity may provide an additional reason for the principal to pay a high wage in the second period only if the stochastic signal indicates that the agent is a reciprocal type. However, we show below that this is also an equilibrium strategy for a selfish principal.

while a selfish type always responds with low effort. Since the agent's type is not observable to the principal when she makes her wage offer, her belief about the share of reciprocal types determines whether she wants to offer the low or the high wage. We impose the assumption that the share of reciprocal agents is not too large, so that offering the high wage is less profitable than offering a low wage—an assumption that is in line with the empirical evidence in MEDIUM.²⁹ Hence, there is again a unique equilibrium in contracting environment MEDIUM, in which wage and effort are always low, in line with what we observe in treatment MEDIUM-LT. Now assume that the historical example about high trustworthiness of the agents in treatment MEDIUM-HT initially renders the principals' beliefs about the share of reciprocal agents more optimistic. The principals may then find it optimal to offer the high wage, in expectation that it will be rewarded with high effort sufficiently often to be profitable. These principals would eventually end up with lower profits than expected according to their beliefs. Learning dynamics that move beliefs gradually towards the truth would then imply that the principals eventually switch to the low wage, exactly in line with the transient effect of the high-trust environment that we observe in contracting environment MEDIUM.

Finally, we model contracting environment STRONG as a repeated sequential-move game. An essential feature of the experimental setting is the finite repetition of the stage game, coupled with an excess supply of agents. This allows principals to rehire or fire agents conditional on the stochastic outcome of their earlier interaction. We capture the dynamic interaction with just two periods. The feature of excess supply of agents is modelled by having one principal and two agents. With repeated interaction, the game exhibits coexistence of a low-trust and a high-trust sequential equilibrium for a large range of parameters, including parameters that resemble the payoff structure in the experiment.

The low-trust equilibrium replicates the outcome of the one-shot interaction. If the principal believes that agents will provide only low effort, she initially offers the low wage to one of the agents, and both types of that agent respond with low effort. The subsequent stochastic realisation of the value is therefore not informative about the agent's type. As a consequence, the principal again offers the low wage in the second period, and both types of agent respond with low effort. A reciprocal agent cannot signal his type by a first-period deviation from equilibrium because the effort choice is not directly observable to the principal. Thus, the low-trust equilibrium is a pooling equilibrium in which gift exchange between the principal and the reciprocal types does not materialise.

In the high-trust equilibrium, the principal initially trusts an agent of unknown type. That is, she pays the high wage. A selfish type responds with low effort, but a reciprocal type responds with high effort. The principal's belief that the agent is a reciprocal type declines if she receives the low value. She will then not offer the high wage again in the second period. A realised high value, by contrast, constitutes a positive signal about the type of agent. Given the positively updated belief, the principal's expected profit is maximised by offering the high wage again to the same agent. A selfish type has no incentive to mimic a reciprocal type because he does not obtain the additional intrinsic benefit from responding to a high wage with high effort. Thus, the high-trust equilibrium is a separating equilibrium, where the initial trusting behaviour of the principal serves to stochastically screen reciprocal types from selfish types.

²⁹ The slope of the relationship between average effort and offered wages is not steep enough in this treatment to generate a positive slope between profits and offered wages. In fact, the relation between average profits and offered wages is negative in MEDIUM.

We interpret the historical information about agents' trustworthiness used in the experiment as a device that selects between these multiple equilibria. The predictions of the high-trust and the low-trust equilibria are indeed in line with the experimental findings of higher wages and higher effort in STRONG-HT compared to STRONG-LT.

More specifically, consider the low-trust equilibrium first. The offered wage and the returned effort are predicted to be low in both periods on the equilibrium path, as confirmed by the experimental results in STRONG-LT. Furthermore, if the principal trembled and offered the high wage in the first period then both types of agent would still respond with low effort in that equilibrium. This is sustained by the correct off-equilibrium belief that the principal subsequently reacts to a realised high value (which is uninformative because both types behave in the same way) by not rehiring the agent. The agent's off-equilibrium behaviour thus confirms the flat wage-effort reaction as induced by the historical example in STRONG-LT.

Consider the high-trust equilibrium next. The high first-period wage elicits an average effort strictly above the low level, as confirmed by our results in STRONG-HT. If the principal trembled and offered the low wage instead, both types of agent would respond with low effort in this equilibrium. This confirms the principal's belief in a positive wage-effort reaction as induced by the historical example in STRONG-HT. A response of effort to wage can also be observed on the equilibrium path of the high-trust equilibrium. Depending on the stochastic realisation of the value, the principal offers either the high or the low wage in the second period, and the induced expected effort is larger in the former case than in the latter.

5. Trust and Contract Enforcement As Substitutes

We have shown that trust and contract enforcement can be complements, over ranges of contractual incompleteness that appear plausible from a real-world perspective. Here, we explicitly demonstrate that trust and contract enforcement can also be substitutes, whenever contracts become sufficiently complete.

Specifically, in our contracting environment labelled TRANSPARENCY, principals observe, not only the performance signals of agents they interacted with, as in our main contracting environments WEAK, MEDIUM and STRONG, but also the performance signals of all other agents in the market. This environment reflects scenarios such as the availability of reference letters from previous employers. Apart from this difference, TRANSPARENCY is identical to STRONG, that is, IDs of all players are fixed and wage offers are binding. The availability of additional information strengthens contract enforcement further because agents are aware that their performance signals are available to all principals, and principals can base their contract offers on the performance signals of agents they have not previously interacted with.

We implemented a single trust environment, LT, in combination with contracting environment TRANSPARENCY. Figure 11 presents the agents' average effort levels over the course of the experiment in TRANSPARENCY-LT and compares them to the average effort levels in STRONG-HT and STRONG-LT. The figure reveals that increasing contract enforcement while holding trust constant—moving from STRONG-LT to TRANSPARENCY-LT—produces the same gains from trade (with effort serving as a sufficient statistic for expected gains from trade) as increasing trust while holding contractual enforcement constant—moving from STRONG-LT to STRONG-HT. The data thus demonstrate that sufficiently high levels of contract enforcement can substitute for low levels of trust.

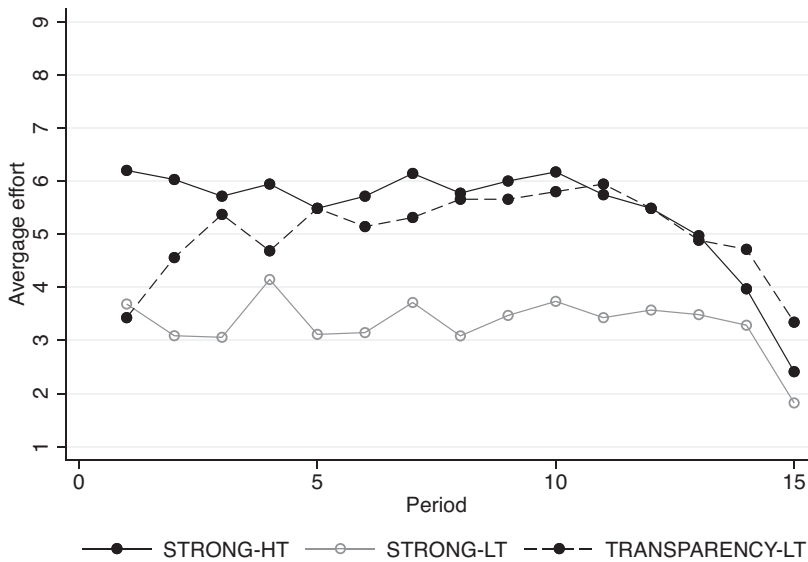


Fig. 11. *Sufficiently Strong Contract Enforcement Can Substitute for High Trust.*

In the limit with perfectly complete contracts, this is trivially the case. In our game-theoretic analysis, which has only a single principal so that we cannot explicitly model observing an agent working for another principal, we model the central feature of contracting environment TRANSPARENCY in reduced form by assuming that providing low effort in response to a high wage offer has an additional reputation cost for an agent. The idea is that information about low performance may spread and later hurt the agent in (unmodelled) interactions with other principals. Under assumptions that are compatible with the assumptions made previously, the modified game has a unique pure-strategy sequential equilibrium outcome whenever the reputation cost is large enough. This outcome involves high wages and high effort levels in the first period and an endgame effect, in line with what we observe in TRANSPARENCY-LT.

6. Conclusions

This paper studies the interplay between trust and the strength of contract enforcement for the realisation of gains from trade by systematically varying both factors independently. It is well understood that informational constraints and weak judicial systems render contract enforcement imperfect. People may then only be willing to interact and realise gains from trade if they trust that their contract partner will not behave opportunistically. By contrast, little appears to be known about how trust affects the causal impact of contract enforcement on gains from trade and how contract enforcement shapes the causal impact of trust.

A better understanding of the interplay of trust and the strength of contract enforcement is important for designing policies aimed at improving economic performance. If trust and contract enforcement were substitutes, policies could be effective if they focused solely on, say, improving the judicial system in order to enable trading partners to better enforce their contracts—even if levels of trust remained low. Policies could be equally effective if they focused solely on,

say, advertising role models of trustful business relations in order to move a society out of a low-trust trap—even if the judicial system remained weak. Our results suggest, however, that such independent improvements along only one dimension—either only trust or only contract enforcement—may not work. We find, in particular, that improvements in the strength of contract enforcement have no or only limited effects on gains from trade when individuals are in our low-trust environment. Likewise, the results indicate that increases in exogenous trust have no or only a transitory impact on gains from trade in our weak and medium contract environments, respectively. In contrast, under strong contract enforcement, trust increases have a large impact, and under high trust, improvements in contract enforcement also have a large effect on the gains from trade. The complementarity between trust and the strength of contract enforcement indicates that simultaneous improvements along both dimensions may be the preferred policy tool.

By documenting empirically that trust and contract enforcement can be complements, and by identifying the mechanisms underlying this complementarity, we also contribute to a deeper understanding of the conditions under which trust and contract enforcement opportunities do and do not exert causal effects on the gains from trade.

As a general implication, research on the determinants of economic performance may benefit by focusing more on interactions between separate factors of influences. Controlling for trust and contract enforcement environments, but not for their interaction, might yield results that obfuscate the real effects. Similar issues may arise for the interaction between other institutional factors, such as legal frameworks or regulatory quality, and informal norms of behaviour, like cultural attitudes towards cooperation or social expectations around corruption. This raises a large range of novel questions for future research.

University of Zurich, Switzerland

University of Zurich, Switzerland

Cornell University, USA

University of Zurich, Switzerland

Additional Supporting Information may be found in the online version of this article:

Online Appendix Replication Package

References

- Aghion, Philippe, Yann Algan, Pierre Cahuc, and Andrei Shleifer. (2010). 'Regulation and distrust', *Quarterly Journal of Economics*, vol. 125(3), pp. 1015–49.
- Alesina, Alberto, and Eliana La Ferrara. (2005). 'Who trusts others?', *Journal of Public Economics*, vol. 82(2), pp. 207–34.
- Algan, Yann, and Pierre Cahuc. (2010). 'Inherited trust and growth', *American Economic Review*, vol. 100(5), pp. 2060–92.
- Arrow, Kenneth J. (1972). 'Gifts and exchanges', *Philosophy and Public Affairs*, vol. 1(4), pp. 343–62.
- Baker, George P., Robert Gibbons, and Kevin J. Murphy. (2002). 'Relational contracts and the theory of the firm', *Quarterly Journal of Economics*, vol. 117(1), pp. 39–84.
- Banerjee, Abhijit V., and Esther Duflo. (2007). 'The economic lives of the poor', *Journal of Economic Perspectives*, vol. 21(1), pp. 141–67.
- Banfield, Edward C. (1958). *The Moral Basis of a Backward Society*, Glencoe, Illinois: Free Press.
- Bartling, Björn, Ernst Fehr, and Klaus M. Schmidt. (2012). 'Screening, competition, and job design: Economic origins of good jobs', *American Economic Review*, vol. 102(2), pp. 834–64.
- Berg, Joyce, John Dickhaut, and Kevin McCabe. (1995). 'Trust, reciprocity, and social history', *Games and Economic Behavior*, vol. 10(1), pp. 122–42.

- Bierbrauer, Felix, and Nick Netzer. (2016). 'Mechanism design and intentions', *Journal of Economic Theory*, vol. 163, pp. 557–603.
- Bloom, Nicholas, Raffaella Sadun, and John Van Reenen. (2012). 'The organization of firms across countries', *Quarterly Journal of Economics*, vol. 127(4), pp. 1663–705.
- Bohnet, Iris, Bruno S. Frey, and Steffen Huck. (2001). 'More order with less law: On contract enforcement, trust, and crowding', *American Political Science Review*, vol. 95(1), pp. 131–44.
- Bohnet, Iris, and Steffen Huck. (2004). 'Repetition and reputation: Implications for trust and trustworthiness when institutions change', *American Economic Review*, vol. 94(2), pp. 362–6.
- Bolton, Patrick, and Mathias Dewatripont. (2005). *Contract Theory*, Cambridge, MA: MIT Press.
- Bower, Anthony G., Steven Garber, and Joel C. Watson. (1996). 'Learning about a population of agents and the evolution of trust and cooperation', *International Journal of Industrial Organization*, vol. 15(2), pp. 165–90.
- Brown, Martin, Armin Falk, and Ernst Fehr. (2004). 'Relational contracts and the nature of market interactions', *Econometrica*, vol. 72(3), pp. 747–80.
- Charness, Ernst. (2004). 'Attribution and reciprocity in an experimental labor market', *Journal of Labor Economics*, vol. 22(3), pp. 665–88.
- Coleman, James S. (1990). *Foundations of Social Theory*, Cambridge, MA: Harvard University Press.
- Danilov, Anastasia, and Dirk Sliwka. (2017). 'Can contracts signal social norms? Experimental evidence', *Management Science*, vol. 63(2), pp. 459–76.
- Djankov, Simeon, Oliver Hart, Caralee McLiesh, and Andrei Shleifer. (2008). 'Debt enforcement around the world', *Journal of Political Economy*, vol. 116(6), pp. 1105–50.
- Ellingsen, Tore, and Magnus Johannesson. (2005). 'Sunk costs and fairness in incomplete bargaining', *Games and Economic Behavior*, vol. 50(2), pp. 155–77.
- Ellingsen, Tore, and Magnus Johannesson. (2008). 'Pride and prejudice: The human side of incentive theory', *American Economic Review*, vol. 98(3), pp. 990–1008.
- Embrey, Matthew, Guillaume R. Frechette, and Sevgi Yuksel. (2018). 'Cooperation in finitely repeated prisoner's dilemma', *Quarterly Journal of Economics*, vol. 133(1), pp. 509–51.
- Falk, Armin, and Michael Kosfeld. (2006). 'The hidden costs of control', *American Economic Review*, vol. 96(5), pp. 1611–30.
- Fehr, Ernst. (2009). 'On the economics and biology of trust', *Journal of the European Economic Association*, vol. 7(2–3), pp. 235–66.
- Fehr, Ernst, Georg Kirchsteiger, and Arno Riedl. (1993). 'Does fairness prevent market clearing—An experimental investigation', *Quarterly Journal of Economics*, vol. 108(2), pp. 437–59.
- Fischbacher, Urs. (2007). 'z-Tree: Zurich toolbox for ready-made economic experiments', *Experimental Economics*, vol. 10(2), pp. 171–8.
- Fukuyama, Francis. (1995). *Trust: The Social Virtues and the Creation of Prosperity*, New York: Free Press.
- Gibbons, Robert. (1998). 'Incentives in organizations', *Journal of Economic Perspectives*, vol. 12(4), pp. 115–32.
- Gibbons, Robert. (2020). 'Visible hands: Governance of value creation – Within firms and beyond', *American Economic Review*, vol. 110, pp. 172–6.
- Gibbons, Robert, Manuel Grieder, Holger Herz, and Christian Zehnder. (2023). 'Building an equilibrium: Rules vs. principles in relational contracts', *Organization Science*, vol. 34(6), pp. 2231–49.
- Gibbons, Robert, and Rebecca Henderson. (2012). 'Relational contracts and organizational capabilities', *Organization Science*, vol. 23(5), pp. 1350–64.
- Greiner, Ben. (2015). 'Subject pool recruitment procedures: Organizing experiments with ORSEE', *Journal of the Economic Science Association*, vol. 1(1), pp. 114–25.
- Guiso, Luigi, Paola Sapienza, and Luigi Zingales. (2009). 'Cultural biases in economic exchange?', *Quarterly Journal of Economics*, vol. 124(3), pp. 1095–131.
- Hart, Oliver. (2009). 'Hold-up, asset ownership, and reference points', *Quarterly Journal of Economics*, vol. 124(1), pp. 267–300.
- Hart, Oliver, and Bengt Holmström. (2010). 'A theory of firm scope', *Quarterly Journal of Economics*, vol. 125(2), pp. 483–513.
- Hart, Oliver, and John Moore. (2008). 'Contracts as reference points', *Quarterly Journal of Economics*, vol. 123(1), pp. 1–48.
- Herweg, Fabian, and Klaus M. Schmidt. (2015). 'Loss aversion and inefficient renegotiation', *Review of Economic Studies*, vol. 82(1), pp. 297–332.
- Knack, Stephen, and Philip Keefer. (1997). 'Does social capital have an economic payoff? A cross-country investigation', *Quarterly Journal of Economics*, vol. 112(4), pp. 1251–88.
- La Porta, Rafael, Florencio López-de-Silanes, Andrei Shleifer, and Robert W. Vishny. (1997). 'Trust in large organizations', *American Economic Review*, vol. 87(2), pp. 333–8.
- La Porta, Rafael, and Andrei Shleifer. (2014). 'Informality and development', *Journal of Economic Perspectives*, vol. 28(3), pp. 109–26.
- MacLeod, W. Bentley. (2007). 'Reputations, relationships and contract enforcement', *Journal of Economic Literature*, vol. 45(3), pp. 595–628.

- MacLeod, W. Bentley, and James M. Malcomson. (1998). 'Motivation and markets', *American Economic Review*, vol. 88(3), pp. 388–411.
- McCaig, Brian, and Nina Pavcnik. (2015). 'Informal employment in a growing and globalizing low-income country', *American Economic Review*, vol. 105(5), pp. 545–50.
- North, Douglass C. (1991). 'Institutions', *Journal of Economic Perspectives*, vol. 5(1), pp. 97–112.
- Putnam, Robert D. (1993). *Making Democracy Work*, Princeton, NJ: Princeton University Press.
- Putnam, Robert D. (2000). *Bowling Alone: The Collapse and Revival of American Community*, New York: Simon and Schuster.
- Sliwka, Dirk. (2007). 'Trust as a signal of a social norm and the hidden costs of incentive schemes', *American Economic Review*, vol. 97(3), pp. 999–1012.
- Sliwka, Dirk, and Peter Werner. (2017). 'Wage increases and the dynamics of reciprocity', *Journal of Labor Economics*, vol. 35(2), pp. 299–344.
- Tabellini, Guido. (2008). 'The scope of cooperation: Values and incentives', *Quarterly Journal of Economics*, vol. 123(3), pp. 905–50.
- Wibral, Matthias. (2015). 'Identity changes and the efficiency of reputation systems', *Experimental Economics*, vol. 18(3), pp. 408–31.