Evidence on Responses to Within-Firm Incentives

a) Within Firms only:
   - Safelite Glass Company (Lazear 2000)
   - “Pay Enough or Don’t Pay at All” (Gneezy and Rustichini 2000)
   - “High Stakes and Big Mistakes” (Ariely et al., 2009)
   - “The Hidden Costs of Control” (Falk and Kosfeld 2006)

b) Implications for Labor Markets:
   - Fehr, Kirchsteiger and Riedl (1993)
   - Brown, Falk and Fehr (2004)
Safelite Glass Co, Columbus Ohio

Largest installer of auto glass in U.S.

Employees work independently; no interactions; output easily measured (quality was handled separately).

Employees were paid on an hourly basis, independent of output. Lazear convinced the CEO to introduce piece rate compensation in mid-1990s.

The new compensation system was introduced on a staggered basis in locations across the country, allowing many before-after comparisons involving 2000 workers.
Predicted effects of introducing piece rates:

Under straight hourly pay, imagine that a worker needed to produce $Q_0$ to keep his job. No financial incentives were provided for additional output beyond $Q_0$. Thus, we expect both ‘lazy’ and ‘ambitious’ workers to produce $Q_0$.

Pay under the new scheme was identical to the old, except that workers who produced more than $Q'$ were rewarded a fixed amount for every additional windshield installed. So we expect ambitious workers, but not lazy workers, to raise their output.

Put a different way, both the mean and the variance of output should rise.
Results:

- variance of output rose; ‘abler’ workers increased their output more

- profits *and* wages both rose

- average output rose by 36%

- the average worker increased his output by about 18%

What accounts for the difference?
Lazear’s data allowed him to:

a) compute average output per worker before and after the change—this increased 36%

b) compute the change in output for those workers who were at Safelite both before and after the change-- this increased only half as much

Thus: half of the increase in output was due to turnover—a different mix of workers before and after the change. The new scheme was more attractive to ‘ambitious’ workers, so:

- they were more likely to stay than ‘lazy workers’, and
- new hires were, on average, abler.

So: Changes in pay policy affect not only existing workers’ effort choices, but the mix of workers attracted to the firm.

HRM innovations can generate massive productivity improvements.
Two experimental tests where subjects were paid by the piece to perform a task. What is the effect of the piece rate on performance?

Focus on the “IQ test”: college students paid per question answered correctly (out of 50).

Subjects all received the same ‘base pay’ (60 NIS show-up fee), then were randomly assigned to be paid piece rates per correct question, ranging from 0 to 3 NIS (just under $1 at the time) per question. Mean performance was:

<table>
<thead>
<tr>
<th>Payment per correct answer (b):</th>
<th>0</th>
<th>0.1</th>
<th>1.0</th>
<th>3.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean performance</td>
<td>28.4</td>
<td>23.1</td>
<td>34.7</td>
<td>34.1</td>
</tr>
</tbody>
</table>

Performance was worse for small rewards than for no reward at all. (similar results from a field experiment involving door-to-door solicitation).

Further: college students do not expect people to behave this way: When student “principals” had to choose between paying 0 and 0.1, 87% chose 0.1!
Possible Explanations:

- Monetary rewards destroy ‘intrinsic motivation”? (Possible, but the effect must be discontinuous).

- “Incomplete contracts”: Offering the reward changes workers’ perception of the labor contract:

  If a zero piece rate is offered, subjects tend to perceive the ‘contract’ as follows: “The experimenter has offered me 60 NIS to do a job. Now I know what that job is—answering questions—and it is my job to do it satisfactorily.” [put a different way—subjects interpret this as a gift-exchange contract]

  If a positive piece rate is offered, subjects tend to perceive the contract more as follows: “The experimenter has offered me 60 NIS to show up. I’ve done that. Now he is offering me 10 cents per question to answer questions. That’s not very much.”
Lesson:

When labor contracts are ‘incomplete’ (which is almost always the case), perceptions and ‘framing’ of what is expected/owed/fair can matter a lot.

They can also be hard to change. See for example:

- Gneezy-Rustichini’s daycare experiment (first introduced, then removed a fine for late pickup). Lateness increased, and never recovered.

- Irlenbusch and Sliwka (IZA working paper no. 1758, Sept. 2005). ‘chosen effort’ lab experiment [no intrinsic motivation likely here]; first they introduce an option for principals to pay a piece rate, then remove this option. Effort fell, then fell even further.
Ariely, Dan; Gneezy, Uri; Loewenstein, George; Mazar, Nina. “Large Stakes and Big Mistakes” *Review of Economic Studies*, vol. 76, no. 2, April 2009, pp. 451-69

In 2009, Dan Ariely and co-authors conducted experiments on how financial incentives affect performance in 6 simple tasks:

-“Packing quarters” (creativity)

-“Simon” (predecessor to “Bop-It”) and “Recall Last Three Digits” (memory)

-“Labyrinth”, “Dart Ball” and “Roll-up” (physical)
The experiments were done in a rural town in India, and compared the effects of “low”, “medium” and “high” incentives on performance.

Low incentives: 4 rupees for “very good” performance
Medium incentives: 40 rupees for “very good” performance
High incentives: 400 rupees for “very good” performance. This is about equal to one month’s earnings in the village.

Their hypotheses:

-going from low to medium incentives might raise performance, but from medium to high might reduce it, due to ‘choking’

-‘choking’ is more likely for creativity and memory tasks, less so for physical tasks.

Their results:

[insert Ariely et al. results here].
FIGURE 1
Means of the share of earnings relative to the maximum possible earnings for the three payment levels. For all six games combined (a), and plotted separately by game (b–d). Games are indicated by their category: motor skills (ms), memory (mm), and creativity (cr).

3.4. Summary

Overall, the results point to three main conclusions: first, with the sole exception of the Labyrinth game there was no significant difference in the performance between the low- and mid-payment conditions. Thus, despite the relative large difference in magnitude of reward across the treatments (i.e. 10 times higher for the mid-payment condition relative to the low-payment condition), performance did not seem to increase. One interpretation of this result is that the incentives in the low payment condition (which were not altogether that low) created a level of performance that was already at a peak.

Second, and more importantly, the performance of participants was always lowest in the high payment condition when compared with the low- and mid-payment conditions together.

Third, and contrary to our expectations, we did not observe any obvious difference in the effect of incentives on performance for different categories of games. We included, for example, “Simon” and “Recall Last Three Digits” because these games require simple memory, and we thought that participants who were more motivated might be more likely to maintain high levels for Packing Quarters, Simon, and Labyrinth, and not significant for Recall Last Three Digits, Dart Ball, and Roll-Up.
Agents are endowed with 120 lab dollars in each round. They choose effort, $x$ (which can be any number, $x$, between 0 and 120), which costs them $x$, but yields a benefit of $2x$ to the principal.

Principals (who move first) only make one decision: whether to impose a minimum effort level, $x_{min}$ on the agent. This restricts the agent’s choice set to the set \{${x_{min}, \ldots 120}$\}. $x_{min}$ was set by the experimenters to be either 5, 10, or 20.
**Predictions:**

1. If the subjects behave like ‘rational economic man’, all agents will choose the minimum effort the principal allows them to, and all principals will decide to restrict their agent’s choice set.

2. Consider an agent who thinks that the ‘fair’ level of effort to provide is $x^* > x_{\text{min}}$. Suppose also that the agent’s perception of what is fair is unaffected by actions taken by the principal that have no consequences for the agent. Then any such agent’s choices should be unaffected by whether or not the principal imposes the requirement that $x > x_{\text{min}}$.

Put a different way, (2) says that the entire empirical distribution of effort levels to the right of $x_{\text{min}}$ should be unaffected by the principal’s decision to impose the minimum effort requirement.

Is it? [insert Figure 1 from Falk-Kosfeld here].
Figure 1. Cumulative Distribution of Agents’ Choices in Treatment C5 (Panel A), C10 (Panel B), and C20 (Panel C)

Notes: The figure shows all observations $x \leq 50$. In each treatment, there were a few $x$-choices above 50. These observations are summarized as $x > 50$. 
Decidedly not: the majority of agents reduces effort when firms attempt to ‘control’ their actions (though the effort levels of a smaller number of ‘opportunistic’ agents’ efforts were mechanically increased by the effort minimum). In most treatments, these net reductions in effort were so substantial that principals who ‘controlled’ earned lower payoffs then those who did not.
**Why did this happen? Here are some clues:**

- When the minimum effort requirement was imposed by the *experimenter*, subjects did not respond by reducing effort.

- In a follow-up survey, the authors asked agents “What do you feel if [the principal] forces you to transfer at least \([x]\) points?” The most common response was ‘distrust’, especially among agents who reacted negatively to control.

- In a clever variation on Falk and Kosfeld’s design, Schnedler and Vadovic (2007) show that control by principals does not elicit negative reactions from agents when the principal’s control is “legitimized” in two alternative ways:

  (a) When the principal must set a common control policy that applies not only to the agent, but also to a computerized ‘automaton’ agent who supplies minimum effort whenever this is allowed.

  (b) When the principal is given a small endowment (in contrast to zero in Falk-Kosfeld), agents do not seem to resent effort requirements that simply prevent agents from ‘pilfering’ this endowment.
Implications:

At least in a laboratory environment, rigid attempts to control agents’ behavior can ‘backfire’. ‘Trusting’ your employees can pay, at least in certain situations.
Effects of Contractual Incompleteness (Moral Hazard) on Labor Markets

Fehr, Kirchsteiger and Riedl “Does Fairness Prevent Market Clearing? An Experimental Investigation” (QJE, 1993)

Four experimental sessions of 12 periods each, all with more “workers” than “firms” (either 9 and 6 or 8 and 5). Workers and firms are in separate rooms.

Each period had two stages: -a one-sided oral auction for ‘jobs’ -exchange of effort for pay

In **stage one** (three minutes): firms offer wages at any time. These are transmitted to workers in the other room, who can accept or reject. Once accepted, that worker and firm are ‘matched’ for the period. Firms can revise their offers if not accepted. Each firm can employ only one worker. Workers (and firms) who remain unmatched get a payoff of zero for the period.

In **stage two**, all matched workers receive their wage, then decide how much effort to supply. This is intended to reflect **contractual incompleteness** in labor markets.
(Employed) workers’ payoffs are:

\[ U = Wage - c - m(e) \]

where \( c = 26 \) for all subjects and sessions) represents the time cost of taking a job, and \( m(e) \) is the disutility of effort. \( m(e) \) is increasing and convex:

<table>
<thead>
<tr>
<th>e</th>
<th>.1</th>
<th>.2</th>
<th>.3</th>
<th>.4</th>
<th>.5</th>
<th>.6</th>
<th>.7</th>
<th>.8</th>
<th>.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>m(e)</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>12</td>
<td>15</td>
<td>18</td>
</tr>
</tbody>
</table>

(Matched) firms’ payoffs are:

\[ \Pi = (v \text{-wage}) \times e \]

Where \( v = 126 \) (for all subjects and sessions). \( ve \) can be thought of as the firm’s revenues; \( wage \times e \) are its labor costs. Strictly speaking, labor costs should not depend on \( e \) when contracts are incomplete. The authors did this to ensure that no firm would earn a negative payoff.
Theoretical Predictions:

Stage 2: regardless of the wage received, agents should pick lowest possible effort: \( e = 1 \)

Stage 1: anticipating this, firms should offer a wage that leaves workers indifferent between employment and unemployment (30).

There will be unemployed (unmatched) workers, but they will be voluntarily unemployed at the equilibrium wage. Both profits and worker utility will be quite low, since little effort is exchanged despite the gains from exchanging more of it.
Experimental Results:

Average wages, at 72, were much higher than the predicted, market clearing level (30).

Average wages, at 0.4, were also much higher than predicted (.1)

Putting these together, firms gave about 42 percent of the surplus generated by the employment relationship to workers (compared to the predicted share of zero).

Both profits and worker utility were much higher than in the market-clearing equilibrium overall surplus was several hundred percent higher).

Unmatched workers were strictly (and a lot) worse off than employed workers: unemployment was *involuntary*.

These results remained essentially constant across all 12 experimental periods (no ‘unravelling’).
Why did this happen—some clues

Workers reciprocated high wage offers with high effort levels.

[show figure I here]

(note that this looks like an upward-sloping labor supply curve, but it’s not….)

This occurred despite complete anonymity: there was no way for any firm to reward any worker for providing high effort. Thus, no incentive for workers to acquire reputation as hard workers.

Firms reported that they tried to induce high effort by offering a high wage.

Workers tried very hard to be the first to accept an offer, and reported that they took the wage offer into account when choosing effort.

Few employers tried to undercut above market-clearing wages by offering lower wages. They usually fared poorly (receiving low effort) and went back to offering high wages.
that were associated with different wages. The pattern of these data points is also clearly sloping upward.

In Table III we present the results of regression (5) with the whole data set as well as with the data of the individual sessions and with the last period observations of all sessions.

Notice that the $\beta$-coefficient is positive and highly significant in all regressions. In Table IV the results of regression (6) are listed.

Table IV confirms the results of Table III with respect to the sign and significance of $\beta$. Furthermore, the adjusted $R^2$s of Table IV are approximately two times the $R^2$s of Table III. As the

<table>
<thead>
<tr>
<th>TABLE III</th>
</tr>
</thead>
<tbody>
<tr>
<td>RESULTS OF REGRESSION (5): $e = \alpha + \beta p + \mu$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$N$</th>
<th>$\alpha$</th>
<th>$t(\alpha)$</th>
<th>$\beta$</th>
<th>$t(\beta)$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1-4</td>
<td>276</td>
<td>-0.18</td>
<td>-3.1</td>
<td>0.0078</td>
<td>9.6</td>
</tr>
<tr>
<td>SL1-4</td>
<td>23</td>
<td>-0.6</td>
<td>-2.2</td>
<td>0.0129</td>
<td>3.5</td>
</tr>
<tr>
<td>S1</td>
<td>72</td>
<td>-0.27</td>
<td>-2.8</td>
<td>0.0076</td>
<td>6.2</td>
</tr>
<tr>
<td>S2</td>
<td>72</td>
<td>-0.34</td>
<td>-2.3</td>
<td>0.0111</td>
<td>5.4</td>
</tr>
<tr>
<td>S3</td>
<td>72</td>
<td>-0.14</td>
<td>-1.6</td>
<td>0.0066</td>
<td>4.9</td>
</tr>
<tr>
<td>S4</td>
<td>60</td>
<td>-0.38</td>
<td>-1.7</td>
<td>0.0113</td>
<td>3.9</td>
</tr>
</tbody>
</table>

$S#$: Session#.
SL1-4: Results of the estimation with the last period data of all sessions.
$N$: number of observations.
t$(\cdot)$: t-value of the relevant coefficients.
$R^2$: Adjusted coefficient of determination.
Lessons:

-when labor contracts are incomplete, ‘internal’ firm concerns (the problem of eliciting effort after the worker has been hired), can affect the nature of labor market equilibrium.

-in this case, the authors have generated a market with involuntary unemployment, where firms who try to undercut above-market-clearing wages lose money.

-this coincides closely with Bewley’s observations that firms are extremely reluctant to cut wages in recessions, due to concerns about the productivity (‘morale’) of the remaining employees.
What happens in these markets if we allow workers to form reputations?

Brown, Falk and Fehr (Econometrica, 2004)

-in a session, ten workers interact with seven firms over fifteen trading periods.

-each trading period had two stages: a market for contracts, followed by the exchange of effort for pay.

-the market for contracts is, again, a one-sided auction: firms posted offers (consisting of a wage, a desired effort level, and the firm’s ID number), which could be accepted or rejected by workers.

-Once a firm’s offer was accepted by a worker, both the firm and worker were removed from the market for that trading period.

-Now, firms can make both private and public offers: private offers were only conveyed to the worker with whom the firm wanted to trade.
BFF compare two main experimental conditions:

Complete contracts (C): the firm’s required effort level was automatically implemented (by the computer) if a worker accepted a contract.

Incomplete contracts (ICF), this third party enforcement of desired effort levels was absent.

In both treatments, firms only observe the effort levels of workers when they are employing them.
BFF’s Findings:

Under complete contracts, markets resemble the textbook case:

- contract offers are public
- long-term relations between firms and workers are absent
- trading parties seem to be indifferent to their partners’ identities
- rent sharing is driven toward the competitive equilibrium (where firms retain all the surplus).

Under incomplete contracts, successful exchange usually proceeds as follows:

- the firm makes a generous contract offer that a worker reciprocates with a high effort level.
- then, the firm repeatedly seeks out the same worker with a private offer.
- rents are shared in these long-term bilateral relationships
- these long-term relationships are disciplined by the threat of non-renewal.
- unemployment is involuntary

Which of these two scenarios strikes you as more representative of ‘real’ labor markets?
In my view, the BFF view of labor markets (where moral hazard problems inside the firm have a profound effect on how the outside labor market works), may be a more useful way to conceptualize labor markets than either the perfectly competitive one, or one of the many equilibrium search models that are so popular nowadays.
Variations on BFF:

**Brown, Falk and Fehr (2008)** replace the excess supply of workers by an excess supply of firms.

Perhaps surprisingly, not much changes: firms pay workers above the going market rate and workers reciprocate with higher effort. Market performance remains high, suggesting that unemployment is not required to enforce high effort levels.

**Falk, Huffman and Macleod (2008)** introduce two institutions -- dismissal barriers and bonus pay.

EPLs lead to large reductions in equilibrium effort and market efficiency. Giving firms the option to reward high effort with bonus pay, however, offsets much of these inefficiencies.