

Field Experiments with Firms

Oriana Bandiera, Iwan Barankay, and Imran Rasul

Firms operate in complex environments: a list of the categories in which they need to make interrelated choices would include employee pay, pricing, product attributes, production technologies, and management. In turn, these decisions involve responding to characteristics that are often hard to measure or uncertain, such as those related to market characteristics, the productivity of individual inputs, and entrepreneurial ability. Due to the complexity of the environment, research that seeks to understand the behavior of firms based on observational data faces many challenges at uncovering causal relationships. In this paper, we illustrate how field experiments, guided by economic theory, can address these challenges and provide new answers to long-standing questions about firms: Do firm choices maximize profits subject to constraints? If so, which constraints bind and inform decision making in firms? If not, why are firms operating inside the frontier?

In this paper, we review field experiments that provide preliminary answers to these questions and map directions for further research. We organize our discussion into two classes of work. The first is field experiments conducted *within* firms, in which the units of observation are workers or divisions of a firm. The theory behind many of these experiments views the firm as an organization, emphasizing agency problems. We discuss field experiments that shed light on solutions to the agency problem, from incentive pay to social pressure and nonmonetary rewards.

■ *Oriana Bandiera is Professor of Economics, London School of Economics and Political Science, London, United Kingdom. Iwan Barankay is Associate Professor of Management, Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania. Imran Rasul is Professor of Economics, University College London, London, United Kingdom. Their e-mail addresses are <o.bandiera@lse.ac.uk>, <barankay@wharton.upenn.edu>, and <i.rasul@ucl.ac.uk>.*

doi=10.1257/jep.25.3.63

The second strand covers field experiments conducted *between* firms, in which the firm is the unit of observation. The theory behind most of these experiments views the firm through the lens of neoclassical production theory, and so we discuss how field experiments have exogenously varied input availability to shed light on constraints firms face.

Throughout, we focus on experiments designed to shed light on firms' behavior. This still leaves out a large class of field experiments that are run in collaboration with firms to provide evidence on other issues such as consumer behavior or optimal auction design.¹

Beyond the results of specific field experiment studies, we also believe that economists can reap enormous benefits from establishing working partnerships with firms and engaging in primary data collection. Thus, we conclude the paper by offering some discussion of the practical issues researchers face in designing and implementing field experiments in firms, and by highlighting research questions that remain relatively untouched by field experiments. We hope that by the end of our discussion, readers have a clear sense of the costs and benefits of field experiments in firm settings, and are motivated to consider this approach themselves.

Field Experiments within Firms

Field experiments within firms are generally designed to shed light on how firms can solve agency problems and motivate their employees. In this section, we review evidence on the two classical solutions to this agency problem—monitoring and pay for performance—as well as more recent work on nonmonetary determinants of motivation such as social relations or status rewards.

Although field experiments within firms have experienced a recent resurgence, they are far from new. One of the first series of field experiments was conducted at the Hawthorne plant of the Western Electric Company, near Chicago, in the 1920s. While the validity of their specific findings has been questioned,² these experiments lay the groundwork for many issues that are now considered part of mainstream personnel economics (Bloom and Van Reenen, 2010a). For example, they led Mayo (1933) to stress that workers are motivated by both monetary and nonmonetary rewards from work, an idea that is being tested by the newest generation of field experiments reviewed below.

¹ Further discussion of field experiments on firms is provided in Levitt and List (2009) where they discuss field experiments related to how consumers respond to product attributes and pricing. On auctions, a nascent literature is now emerging that uses field experiments to measure, for example, reserve price effects (Reiley, 2006; Brown and Morgan, 2009; Ostrovsky and Schwarz, 2009).

² As one example, between 1924 and 1927 the level of lighting was systematically changed for experimental groups in different departments (Mayo, 1933). Levitt and List (2011) recently recovered the thought-to-be-lost data from this experiment, and find little evidence that workers reacted to the differences in lighting.

Monitoring and Shirking

The standard agency framework with asymmetric information views employees as rational shirkers: that is, employees consider the marginal costs and marginal benefits of shirking, and decide on their level of effort. Firms thus choose compensation and monitoring policies with shirking in mind. The theory suggests that a reduction in monitoring will tend to an increase in shirking. Moreover, an increase in shirking resulting from reduced monitoring should be greatest among individuals for whom the ongoing employment relationship is least valuable. Three concerns have plagued nonexperimental approaches to testing these ideas: 1) shirking behavior is hard to detect; 2) the ability of the econometrician to detect shirkers might be endogenously related to the employer's monitoring practices; and 3) unobserved factors, such as hiring policies, may lead monitoring and shirking outcomes to be correlated.

Nagin, Rebitzer, Sanders, and Taylor (2002) address these challenges using a field experiment run by a telephone solicitation firm across four of its 16 call centers. At each call center, telephone solicitors were paid a piece rate in which salary increased with the number of "successful" solicitations—where success was reported by the employees themselves. This piece rate created incentives for employees to overstate whether a donation had been promised. To curb opportunistic behavior, the employer monitored by calling back a fraction of those who were reported to have responded positively to a solicitation. Employees were informed when hired that their activities would be checked by callbacks. The results of each week's callbacks were communicated to both employees and their immediate supervisors, and calls found by the monitors to be unsuccessful were deducted from each individual's weekly incentive pay. Stronger sanctions for such calls were not generally imposed on employees because it was understood that donors sometimes change their minds after agreeing to pledge money.

To see if the costs of this monitoring system could be reduced, the company simulated a lower audit rate by experimentally varying the fraction of bad calls that were reported back to employees and supervisors in the four experimental sites from 0 to 2 to 5 to 10 percent, while keeping the actual audit rate at 25 percent in all four sites. By working with the firm, the researchers were able to collect survey data on employee attitudes toward the job, their expected job tenure, and their perceived difficulty of finding a comparable job.

The findings indicate that workers' responses are very heterogeneous. Between 10 and 41 percent of the employees in the four experimental sites behave as "rational cheaters"—that is, they respond to a reduction in the perceived cost of opportunistic behavior by increasing the rate at which they shirk. The remaining 59 to 90 percent of employees, however, do not increase shirking following the reduction in monitoring rates. Using the survey data collected, the authors find that those employees who responded to reductions in monitoring tended to be those who perceived the employer as being unfair and uncaring and that, in contrast with the rational cheater model, individuals with good outside options did not increase shirking by more than other workers when the rate of monitoring declined. This

heterogeneity implies that the optimal monitoring scheme will need to balance the requirement to reduce the shirking behavior of some workers inclined to rationally cheat, against the desire to avoid monitoring costs for those unlikely to cheat under normal circumstances.

Monetary Incentives

In the many circumstances in which monitoring is not practically feasible, the agency problem can be addressed by designing incentive schemes that seek to align the employees' interests with the principal's. A wide class of schemes such as piece rates, bonuses, and prizes achieve this goal by making the employees' pay an increasing function of their performance.

Agency theory makes precise that such "pay for performance" schemes affect productivity both by increasing the productivity of existing employees (the incentive effect) and by attracting more-productive employees to the firm (the selection effect). The incentive effect arises because pay-for-performance schemes increase the marginal benefit of effort, which leads employees to work harder, other things equal. The selection effect arises because high-ability employees who are capable of achieving high performance can achieve higher pay and are, thus, attracted by schemes that reward performance, other things equal. Both the incentive and selection effects increase the variance as well as the mean of productivity and pay, as more-able workers can respond more to the increase in incentive power.

Testing the rich predictions of agency theory using nonexperimental data faces serious econometric challenges, most importantly that observed incentive contracts might be endogenous to firm performance (Prendergast, 1999; Chiappori and Salanié, 2003). Field experimenters can tackle this challenge directly by engineering exogenous changes in incentive schemes across, or most commonly, within firms. Moreover, high-powered incentive schemes might affect productivity through channels that are typically not measured in secondary data. For instance, the increase in pay inequality can reduce workers' morale and lead to sabotage, or the change in the composition of the peer group due to the selection effect might affect workers' behavior over and above the effect of the compensation scheme. Field experimenters are well placed to work with firms to collect primary data on the relevance of these mechanisms, for instance by measuring social ties within the firm.

Among the first of the field experiments designed to measure incentive effects of monetary compensation schemes is that of Shearer (2004), who estimates the productivity gains moving from a fixed wage to a piece rate scheme for tree planters in British Columbia, Canada. Workers were randomly assigned to plant under one of the incentive schemes at the start of a work day. Workers' productivity increased by 20 percent moving from fixed wages to piece rates. In line with the prediction of agency theory, the standard deviation of output across workers was higher under piece rates. Shearer also develops and estimates a structural model of workers' behavior to shed light on: 1) what would have been the productivity gains if management had been imperfectly informed about planting conditions; and 2) how workers would have responded to an efficiency wage scheme. This is a

“natural field experiment,” to use the terminology explained in the introduction to this symposium and in Harrison and List (2004), because the workers did not know that the experiment was occurring.

Combining research methods like this is to be encouraged. In many scenarios, it would be ideal to combine evidence from field experiments with structural modeling to posit an underlying behavioral mechanism behind the effects, to assess the sensitivity of the estimates to slight alterations in the economic environment, and to make headway in understanding the optimal compensation structures. Of course, the validity of the structural model can itself be tested by exploring whether it predicts the responses observed to the exogenous variation engineered by the field experiment. For example, Cho and Rust (2010) follow this approach in using a field experiment to validate a structural model of rental rates for automobiles.

While recent field experiments have made substantial progress in identifying the causal effect of performance pay on workers' effort, they have been less successful at pinning down selection effects.³ This is a priority for future research as the available nonexperimental evidence suggests that selection effects are at least as important as incentive effects. In a nonexperimental study that exploits the roll-out of a new piece rate scheme in a manufacturing firm, Lazear (2000) shows that selection effects explain half of the 44 percent increase in worker productivity that followed the introduction of piece rates. However, identifying selection effects poses a difficult challenge for field experimenters both because it requires information on the entire pool of potential employees and because the time horizon of field experiments is often considerably shorter than that needed for existing workers to quit and new workers to join a firm. Varying incentives across divisions or plants of the same firm while allowing employees to move across divisions might be a way to address both issues. Greater knowledge about selection effects would also help in understanding whether and how the compensation policies of a given firm have spillover effects on other firms that compete for similar workers—which in turn would help tie together the two disparate literatures on within-firm compensation policies and equilibrium wage-setting behavior.

Monetary Incentives and the Social Organization of the Workplace

Until recently, the importance of the interaction between social relations and monetary incentives in the workplace has been addressed mainly in the organizational and business sociology literatures. However, such concerns have begun to be incorporated in economic theory; for example, Kandel and Lazear (1992) and Rotemberg (1994) extend standard agency models to take into account peer effects and social concerns. Moreover, credible evidence on the existence and magnitude of such social mechanisms has begun to emerge using nonexperimental methods in combination with personnel data (Mas and Moretti, 2009; Bandiera, Barankay, and Rasul, 2010). Identifying the causal effects of social relations using nonexperimental

³ Niederle and Vesterlund (2007) explore selection effects of incentives in a laboratory setting.

data faces two main challenges: 1) the observed variation in incentives might be correlated with other unobservable determinants of performance; and 2) firm personnel records rarely contain information on social connections within the firm.⁴

In this section, we discuss a series of three of our own field experiments (Bandiera, Barankay, and Rasul, 2005, 2007, 2009, 2011) that provide novel evidence on the interplay of incentives and the social organization of the workplace, namely the social relations that exist between a group of coworkers or between workers and managers. The firm we study is a leading U.K. producer of “soft fruit”—a broad category that might include fruits like plums, cherries, peaches, strawberries, raspberries, and grapes. The firm’s hierarchy has four layers: the owner and chief executive officer, the general manager, field managers, and workers. The main task of the bottom-tier workers is to pick fruit. Around 40 workers pick on any given field on a given day. Within a field, workers are allocated their own row of fruit to pick, and worker’s productivity is defined as kilograms picked per hour. Managers organize field logistics: for example, they assign workers to rows and make sure workers’ full crates are replaced with empty ones. Managers also choose how much effort to exert and how to allocate this effort among different workers. In this setting, managerial effort is complementary to worker’s effort; for example, if a manager reassigns a worker to a new row as soon as that worker is done with the previous one and removes that worker’s full crates quickly, then even for a given effort level of the worker, the worker will be more productive.

The general manager, who is a permanent employee of the farm, decides which of the workers present on the farm are selected to pick fruit each day, which are assigned to non-picking tasks, and which are left unemployed for the day.

In our setting, workers and managers are hired seasonally from Eastern Europe and live on the farm for the duration of their stay; thus, they have opportunities to form social connections. These connections can be measured by asking workers to report colleagues to whom they are socially linked, or indirectly by using common characteristics—like a common language—that predict social links. In two of the three experimental seasons, the group of coworkers that a given worker is assigned to work with varies across fields and days, and this variation is orthogonal to other determinants of productivity. This creates plausibly exogenous variation in the presence of socially connected workers that can be used to identify the effect of social connections on behavior and productivity.

Our three experiments ran during the 2002, 2003, and 2005 picking seasons. The workforce changes annually, so that workers and managers are exposed to one experiment only. Each experiment induces exogenously timed changes to the incentive structure of one layer of the hierarchy. In all cases the experimental treatments are applied simultaneously to all relevant agents. The rationale behind the within-subject design is that, like most other firm settings, it was impossible to prevent information spillovers between treated and control groups. Moreover, the

⁴List and Rasul (2011) provide a more comprehensive review of this literature.

composition of the workforce is fixed for the duration of each season, thus we focus on incentive rather than selection effects throughout. One aspect of the experimental design is that time-varying unobservables represent the main identification threat. This is addressed by allowing flexible interactions with time effects and by using difference-in-differences estimators that combine data from the experimental seasons and from 2004, during which no experiments were implemented.

In our first experiment, we exogenously varied the workers' compensation scheme from "relative compensation" to piece rates. Under relative compensation, workers are paid a unit price for each kilogram of fruit picked that is negatively related to average productivity on the field-day—thus, if average productivity on a certain day was high, the unit price paid to workers is lower. Under piece rates, workers are paid for each kilogram of fruit picked. Under relative compensation, each worker imposes a negative externality on colleagues: any worker who increases effort will raise average productivity and so reduce the unit wage for all coworkers on the field. Under piece rates, this externality does not exist. As a consequence, under relative compensation, the socially optimal level of effort is lower than the private optimum, whereas under piece rates, the two coincide. The comparison of productivity under the two schemes reveals whether, and to what extent, workers are able to cooperate.

In our 2005 paper, we show that productivity is 50 percent higher under piece rates. Calibration of the maximization conditions of worker's individual effort choice problem reveals this productivity differential to be consistent with the assumption that under relative incentives, workers internalize the negative externality their effort imposes on coworkers. This social incentive is equivalent to them placing a weight of two-thirds on all coworkers' pay. We find that workers internalize the externality more when they work alongside their friends as opposed to colleagues to whom they are not socially linked, and this effect is larger in smaller groups. Yet, the effects disappear under piece rates. Finally, we find that cooperation collapses when workers cannot monitor each other; specifically, moving to piece rates does not increase workers' productivity when they pick from a plant whose physical characteristics make it difficult for a worker to see colleagues. This finding rules out pure altruism as a mechanism to sustain cooperation in this setting.

Our second experiment exogenously varied managers' pay from fixed wages to fixed wages plus a performance bonus that increases in the average productivity of the workers managed. In this experiment, workers were paid piece rates throughout. Our 2007 paper shows that the introduction of managerial bonuses increases both the mean and the dispersion of workers' productivity. One reason, as theory suggests, is that after the introduction of performance pay, managers targeted their effort towards more-able workers. Another reason, again suggested by theory, is that workers with the highest productivity were more likely to be selected into the workforce when managers were paid performance bonuses. Least able workers were employed less often and workers at the bottom of the productivity distribution were never selected to pick.

Our 2009 paper further analyzed the data from this second experiment, with a focus on the interplay between managerial incentives and the social connections between workers and managers. We find evidence that when managers were paid fixed wages, they targeted workers to whom they were socially connected, regardless of ability level. When managers were paid as a function of firm performance, they targeted high-ability workers, regardless of social connections. The findings suggest that social connections have a sizable impact on productivity: when managers were paid fixed wages, the average worker was 9 percent more productive on days when they were managed by someone to whom they were socially connected.

Our third experiment evaluates different compensation schemes under team production. From 2005 onwards, workers were organized in teams of five. Workers were allowed to choose the composition of their team as long as all five members agreed. Compared to previous seasons, the social organization of the workplace is thus endogenous. In Bandiera, Barankay, and Rasul (2011), we compare three forms of team incentives: team piece rates, rank incentives (namely publicly provided information on each team's productivity), and monetary prizes assigned to the most productive teams. The experiment is again closely tied to an underlying model that suggests two key forces that drive team formation: workers' ability and social connections. On one hand, workers have incentives to match by ability. On the other hand, workers might prefer to form teams with friends, both because this might limit free-riding within teams and also because they enjoy nonpecuniary benefits from interacting with coworkers to whom they are socially connected. Our experiment is designed to exogenously alter the trade-off to sorting into teams by friendship relative to ability.

We show that strengthening incentives, either through rank incentives or monetary prizes, makes workers more likely to form teams with others of similar ability instead of with their friends. However, rank incentives and monetary prizes have opposite effects on average productivity: rank incentives significantly reduce it by 14 percent, while monetary prizes significantly increase it by 24 percent. Both effects are heterogeneous: rank incentives only reduce the productivity of teams at the bottom of the productivity distribution, and monetary prizes only increase the productivity of teams at the top. Focusing on the teams that remain intact after each change in incentives, we show that the documented negative effect of rank incentives is primarily due to the endogenous changes in team composition, rather than changes in behavior of the same team. In contrast, the provision of monetary prizes affects firm performance through both the endogenous changes in team composition and changes in behavior within the same team.

Taken together, this set of field experiments yields several lessons. Social connections can drive behavior in the workplace: workers and managers internalize the effect of their effort on colleagues to whom they are socially connected. Moreover, social connections and monetary incentives interact, and the extent to which agents internalize the externality depends on the strength of monetary incentives. Firms should take these potential interactions into account. Indeed, such differences in the social organization of workplaces might help explain some part of

the productivity differences in otherwise observationally similar firms. This set of experiments also illustrates the advantages of combining the variation exogenously created by the field experimenter (the incentive scheme) with other sources of variation that occur naturally in a real-world environment (social ties, monitoring possibilities, peer groups). This combination of experimental rigor and collecting primary data is perhaps the most attractive feature of field experiments.

The findings open up new questions for theoretical research on organizations. For instance, relative incentives led to lower productivity because workers internalized the negative externality to some extent. This finding speaks directly to Lazear's (1989) observation that workers are rarely compensated according to rank-order tournaments, and it points to new and interesting directions for the development of theory concerning the optimal provision of incentives under more-robust assumptions on worker preferences. More broadly, the findings raise the issue of whether incentive policies are indeed chosen optimally, or whether firms are effectively within the efficiency frontier. As we shall see, other field experiments also cast doubt on whether firms make optimal choices. The concluding section will bring these together and discuss implications for future research.

New Topics in Within-Firm Field Experiments

A vigorous literature based on within-firm field experiments is beginning to emerge. Some of the topics focus on organizational features of firms, extending the kind of approaches discussed above. For example, new field experiments are being designed to test alternative ways to motivate employees. A natural candidate is nonmonetary incentives in the form of status or social recognition rewards, such as "employee of the month" job titles. The notion that individuals crave status has long been studied and more recently formalized (Moldovanu, Sela, and Shi, 2007; Besley and Ghatak, 2008). In a field experiment run in collaboration with a public health organization, Ashraf, Bandiera, and Jack (2011) randomize 800 community agents hired to sell condoms in urban compounds into four monetary and nonmonetary rewards treatments. Agents who are assigned to the nonmonetary rewards treatment—namely, stars for performance plus a public ceremony for top performers—sell twice as many condoms as agents who are offered a financial margin on each pack sold.

Another personnel policy that is being subject to experimental scrutiny is the provision of performance feedback. In a recent field experiment conducted with 330 employees recruited via Mechanical Turk, a platform run by Amazon.com for work submitted online, Barankay (2010) finds that the provision of individual performance feedback about relative performance reduces the productivity of workers.

Despite rapid progress in these areas, evidence on other key organizational features is still lagging behind theory. The best example is perhaps the distribution of authority within firms, which has been at the core of theoretical studies of the firms since Coase's (1937) seminal contribution. Modern theoretical work highlights the role of authority as a determinant of incentives (Aghion and Tirole, 1997)

or as a coordination device (Garicano, 2000; Alonso, Dessein, and Matouschek, 2008). We envisage future field experiments designed to shed light on the effect of organizational design on firm performance.⁵

Field Experiments across Firms

Field experiments that take the firm as the unit of observation often seek to exogenously vary the availability of key inputs and in this way seek to uncover the constraints faced by firms. Many of the experiments reviewed in this section are implemented in developing countries from South Asia to Latin America, both because identifying the constraints faced by these firms is key to understanding the development process and because, in practical terms, it is cheaper to create sizable shocks to inputs when the value of a firms' stock of inputs is small.

Physical Capital and Access to Finance

A field experiment by de Mel, McKenzie, and Woodruff (2008) among small and medium enterprises in Sri Lanka illustrates how field experiments can illuminate the long-standing question of how credit market imperfections and liquidity constraints may affect firm growth. They sample 408 enterprises, equally split between retail sales and manufacturing/services, that have less than \$1,000 invested in capital. Around half were randomly assigned to receive one of four treatments: \$100 in cash, \$200 in cash, \$100 in inventories or equipment, or \$200 in inventories or equipment—in the last two treatments, as chosen by the firm owner. These transfers are large compared to the existing capital stock and median monthly profits. The experiment was accompanied by a quarterly panel survey on investment decisions, profits, and also personal characteristics of the owner such as wealth, risk aversion, and cognitive ability. The experiment was framed as random compensation for participation in the survey. Hence owners did not know about the existence or scope of the experiment.

The experiment yields three key findings. First, both types of transfers increase capital stock and profits. Transfers also increase the hours worked by the owner, indicating complementarity between capital and labor hours. Second, the return to the additional capital is around 5–6 percent per month, giving a real annual return well above the market lending rate. Third, the authors develop a model that makes

⁵ Nonexperimental evidence on the determinants of the distribution of authority and its effects on firm performance is also being established. Bloom, Sadun, and Van Reenen (2010) develop a survey instrument to measure the level of centralization of decision making in 4,000 firms across 12 countries, which shows that highly centralized structures are more likely to occur in developing countries and where product market competition is low. Wu (2011) exploits a natural experiment and detailed personnel data from a Chinese newspaper to provide evidence on the effect of centralizing decision-making authority on the effort and performance of managerial editors and reporters. In his setting, centralizing authority reduces the effort of managers and increases the effort of workers, overall increasing the quality of the average article. More importantly, authority has a significant impact on incentives despite the fact that reporters were already on high-powered performance pay.

precise how missing credit or insurance markets can generate the observed discrepancy between returns to capital and lending rates. The balance of evidence indicates that results are driven by missing credit markets, not by risk aversion to borrowing.

Several methodological points are of note. First, using GPS coordinates, the authors show that the treatment has spillover effects on nearby firms. Interestingly, the authors show that spillovers are entirely driven by firms in the bamboo industry, where the harvesting of bamboo is subject to government restrictions and treated firms crowd out others by purchasing all of the available supply. These findings illustrate the potential of field experiments to shed light on the functioning of markets, not just firms in isolation. Second, the authors show how to deal with an attrition rate that was 5 percent higher for control firms than for treatment firms by using the methodology proposed by Lee (2009), which estimates upper and lower bounds for the treatment effect and allows them to show that the estimated treatment effects and return to capital are robust to attrition. Third, the authors also compare the difference between experimental and nonexperimental methods. Compared to a 5.3 percent monthly return to capital estimated via experimental methods, ordinary least squares, random effects, and firm fixed effects models yield estimates of 2.6, 1.7, and 0.07 percent, respectively.

McKenzie and Woodruff (2008) implemented the same experiment with 137 small and medium-sized retail firms in Mexico, and also find high rates of returns to capital, especially for firms that report being credit constrained/not having access to formal loans. Overall, these field experiments provide a useful answer to an important question, as well as providing guidance for future field experiments. In particular they illustrate how experiments can go beyond understanding the decision process of a single firm and provide evidence on the nature of spillovers and market interactions.

Managerial Capital and Enterprise Training

Economic theory has long taught that managerial capital is an important input in production (Lucas, 1978; Rosen, 1982). However, systematic empirical evidence about the effects of managerial capital was, until recently, almost nonexistent—in part because of the difficulties in measuring managerial capital. In addition, managerial capital seems likely to be related to characteristics of the firm and perhaps to unobservable traits about management and workers in ways that made it very difficult to draw causal inferences.

Two recent methodological improvements have led to substantial progress in this area. First, some studies have focused on management practices that can be measured systematically. As they discussed in this journal, Bloom and Van Reenen (2010b) have developed a standard measure of management practices and used it to survey managers in 6,000 manufacturing firms in 17 developed and developing countries. Their work shows a robust correlation between the quality of management practices and firm performance both across and within countries. Second, researchers have begun designing field experiments aimed at evaluating the causal impact of managerial practices on firm performance.

To evaluate the effect of business training on the performance of microenterprises, Karlan and Valdivia (2011) and Drexler, Fischer, and Schoar (2011) examine the experience of clients of a microfinance institution in Peru and the Dominican Republic, respectively. Both studies randomize microentrepreneurs into a treatment group that receives financial training, including basic accounting, marketing, and pricing, and into a control group that does not. Common findings emerge from these studies. Neither finds an effect of business training on sales, profits, or employment. Both find that training reduces business seasonality by increasing sales in “bad” months. These are intriguing results and should lead to significant advances in understanding the constraints faced by microenterprises.

Bloom, Eifert, Mahajan, McKenzie, and Roberts (2011) develop a field experiment to evaluate the effect of modern management practices on the performance of large Indian firms in the textile industry. Working in collaboration with a leading international management consulting firm, the researchers offered free management consulting to a randomly selected group of 14 out of 20 plants belonging to 17 large manufacturers. A further eight plants belonging to the same firms were also surveyed. The consulting intervention targeted 38 key practices that capture standard manufacturing principles in high-income countries.

The design of this field experiment illustrates the trade-off between sample size and the complexity of the intervention needed to study a cross-section of large firms. Consultancy and data collection costs—\$75,000 per treated firm, \$20,000 per control plant—limited the sample size to 20 units. Standard statistical tests that rely on asymptotic properties cannot be used in this context because the number of observations is too small. Collecting data over a long time horizon partially helps as there are procedures that rely on asymptotic approximations along the time dimension (Ibragimov and Müller, 2010). In addition, statistical power can be sustained in such small samples by collecting data directly from machine logs, focusing on similar firms using identical technologies, and collecting high-frequency repeated measures (McKenzie, 2011). The key finding is that managerial capital improved quality and efficiency, reduced inventory, and raised average productivity by 11 percent. The resulting increase in yearly profits was estimated to be over 90 percent of the market cost of the consulting services that firms would have paid in the first year.

Again, such findings raise the obvious questions of why profitable practices were not adopted before the intervention. Interviews with owners and senior managers reveal that incorrect beliefs about the profitability of the practices were the main cause of non-adoption, but even when these incorrect beliefs were pointed out, firms were very slow to adopt the new practices. Most of the owners attributed this to lack of time. However, it also seemed that most Indian textile firms did not need to adopt the new practices in order to compete with other domestic firms as high tariffs shielded them from international competition. The study thus raises what seems to be a recurrent question in field experiments with firms: What are the constraints that discourage firms from making changes that clearly seem to be profitable?

Labor

An alternative kind of between-firms study, more common in high-income countries, are field studies designed to measure discriminatory practices by exogenously varying the applicant pool available to firms. The typical “audit” study presents employers with two sets of job applicants who are identical along all relevant employment characteristics except the characteristic of interest, like race, gender, or age (Heckman and Siegelman, 1993; Riach and Rich, 2002).

Bertrand and Mullainathan (2004) design a field experiment along these lines by sending résumés with randomly assigned white- or black-sounding names to over 1,300 help-wanted ads in Boston and Chicago newspapers. They also randomly vary the quality of the résumé by adding experience, skills, or honors. They find that white names receive 50 percent more callbacks for an interview than black names. This racial gap is uniform across occupation, industry, and employer size. Additionally, the return to higher-quality résumés is higher for whites than blacks, implying that the racial gap is larger for more-qualified applicants.

The authors findings are consistent with a model of lexicographic search whereby the employers stop reading (and hence fail to see all credentials) once they see a black name. While such practice is indeed consistent with the findings, more work is needed to provide direct evidence on its relevance compared to statistical and taste-based discrimination models. If the search process is lexicographic, future work needs to understand why such methods emerged in the first place. Time constraints are one possible ingredient here; indeed, although time constraints are rarely made explicit in economic modeling of firms, they may help to explain experimental results in many diverse contexts. More generally, these types of audit experiments could be fruitfully applied to analyze the practical relevance of other applicant traits such as gender, education, and past employment history. This would allow one to provide a consistent picture of firms’ preferences and constraints on hiring decisions.

Practical Considerations

The growing literature using field experiments on firms has begun to provide insights on long-standing areas of economic research related to firm behavior. In the hope that some readers may be interested in undertaking this kind of research, in this section we discuss two practical considerations that arise in doing such experiments: design and ethical issues. In the final discussion, we then draw together some common lessons from the studies presented to highlight some areas to which we think future research should be directed.

Design

The design of a field experiment ought to be grounded in economic theory so that null and alternative hypotheses can be mapped back to an underlying model of firm behavior. Once null and alternative hypotheses are well defined, this will guide the collection of primary and secondary data.

Researchers then need to choose the unit over which to introduce experimental variation: for field experiments involving firms, this can involve engineering experimental variation across firms or within a firm. For the former, the unit might be inputs, whose characteristics are exogenously varied across firms, or the unit might be firms themselves who are then exogenously assigned to different environments. This latter type of field experiment design remains scarce, but we discuss potential developments below. For experimentation within firms, the unit is most often workers, although plants or firm divisions might also be used. For experimentation at the worker level, at one end of the spectrum, workers are simultaneously randomly assigned to control and treatment groups, as is typically done in policy evaluation and in randomized controlled trials. At the other end of the spectrum, all workers are treated but the timing of the treatment is exogenously chosen by the researchers. Each approach has its own costs and benefits, and these are likely to differ across firm settings.

The main benefit of having a control group is that common trends can be weeded out by using a difference-in-differences estimator. The main cost is that the estimated effect of the intervention might be biased because the control group might react to *not* having received the treatment. Whether this creates a positive or negative bias depends on whether the control group tries to differentiate themselves from, or to emulate, the treated group. This is a first-order issue in field experiments with firms, where it is often harder to isolate treatment and control groups geographically or informationally, so those in a control group are likely to find out about the experiments. The threat of contamination can be eliminated by separating the groups, but this typically causes them to be subject to different workplace conditions, making the control group a weaker counterfactual for what would have happened to the treated in the absence of treatment.

Switching all agents between control and treatment groups at an exogenously chosen time has the benefit of eliminating the contamination bias; also, to increase statistical power, the effect of treatment can be estimated by comparing each agent to that same agent without the treatment, thus eliminating all sources of unobservable heterogeneity (List, Sadoff, and Wagner, 2010). The cost of this approach is that the estimated effect of the treatment might be biased because of unobservable determinants of changes in behavior. This concern might be addressed by collecting a sufficiently long time series during both treatment and control periods, or, if seasonality or cyclicalities is a potential threat, by collecting information from a different period during which no treatment was introduced, to purge estimates of variation due to such naturally occurring fluctuations. Switching all agents between control and treatment groups is a good approach for another, practical reason: firms often express unwillingness to treat similar workers in the same plant or firm location in different ways.

A recent field experiment illustrates how different randomization strategies can lead to different estimates. Shi (2010) compares productivity under fixed wages and piece rates for workers engaged in tree thinning in a fruit orchard in the state of Washington. As the firm has multiple sites, in one site she switched ten workers

simultaneously from wages to piece rates and observed them for three days under each treatment, whereas in another site she switched seven workers to piece rates and kept another seven as controls with fixed wages for one day, identifying the effect from the difference-in-differences between the two days. The estimated productivity increase is 23 percent in the first design—similar to the estimates in Shearer (2004)—while the estimated effect increases to 43 percent in the second design. Shi (2010) reports that workers in the control group learned of the existence of a treatment group and were not pleased. This could have reduced their productivity, leading to an overestimate of the effect, but her data is not rich enough to shed light on the mechanism. Understanding such contamination effects is crucial to being able to compare findings across experiments, and such issues are beginning to be explored by field experimenters.

Ethics

Field experiments involve human subjects and thus typically fall under the oversight of an institutional review board run by an institution of higher education or a funding agency. In turn, these boards are typically guided by the ethical principles set forth by a U.S. government report called the Belmont Report (named after the conference center where it was drafted in 1978). The three fundamental ethical principles in the Belmont Report for all human subjects research are respect for persons, beneficence, and justice.⁶ Recent years have seen some convergence in institutional review board practices across institutions.

One question of particular interest to economists is whether all participants in an experiment must consent in advance, as the knowledge of participating in an experiment may bias the results. A common rule is that institutional review boards may waive informed consent and allow the use of such “deception” *if* certain conditions are met: First, the research involves no more than minimal risk. Second, the waiver will not impact adversely subjects’ rights and welfare, which includes no reduction in compensation, employment benefits, or mental well-being. (Moreover, if subjects do not wish to participate in the research there is to be no adverse effect on them either.) Third, the research could not practically be carried out without the waiver. Finally, subjects will be provided with additional pertinent information about participation—a condition which can often be met by debriefing subjects at the end of the study.

The granting of waivers of informed consent varies across institutional review boards. Some simply allow the waiver. Other boards require subjects to be informed that they are part of a research study (rather than an “experiment”), that they can opt out of the study without any consequence, and that they are provided with the contact details of the researcher. This need not compromise the conduct of field experiments within firms as long as such requirements are equally applied to treatment and control groups.

⁶The report (National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979) is located at (<http://www.hhs.gov/ohrp/humansubjects/guidance/>).

A distinctive characteristic of field experiments with firms that has important implications for how field experiments should be judged by institutional review boards is that firms are likely to experiment on their own, or be advised by for-profit consultancies, and the ethical guidelines they must adhere to in such cases are likely to be much less stringent than those faced by academics.⁷ Academic researchers might then be crowded out by for-profit evaluation consultancies that are not subject to the same ethical requirements, and this might reduce the involvement of academics in field experiments and create a strong selection among the type of experiments that can be subject to scientific scrutiny. The severity of this concern will depend on the evolving attitude of institutional review boards with regard to academics' involvement with firms. For instance, will researchers be allowed to advise firms on experiments that do not meet academic ethical guidelines but would be implemented regardless of academic involvement? Will researchers be allowed to analyze data from experiments that do not meet academic ethical guidelines but were initiated by the firms themselves? The answers to these questions will have profound implications for the future of field experiments with firms.

Common Lessons and Future Directions

Field experiments are at the heart of a growing empirical literature that is expanding economists' understanding of firm behavior. In this concluding discussion, we draw together some common lessons from these studies and suggest some future directions.

A common feature of most field experiments with firms is that they identify partial equilibrium effects, as they typically affect only one or at most a sample of firms within one industry. General equilibrium effects might of course differ from the partial equilibrium effects, depending in part on whether the returns to the experimental innovation can be competed away. For instance, if high-powered incentives benefit the firm by attracting better workers, no firm can gain by offering high-powered incentives when all firms do the same. In contrast, if high-powered incentives increase workers' productivity, then all firms can benefit from offering high-powered incentives, thus increasing aggregate productivity.

The question that then naturally arises is: if these aggregate gains are possible, why don't firms reap them? Indeed, one puzzling finding is that almost all the field experiments reviewed have brought large benefits to the firm. In part, this pattern arises because firms would not agree to implement experiments that are expected to have detrimental consequences! Yet the fact that in so many cases researchers have managed to increase profits appears at odds with the common assumption that firms are pressured by competitive forces to make at least close-to-optimal choices. For example, the incentive schemes introduced in our field experiments

⁷We thank the editor, David Autor, for raising this point.

at the fruit-picking firm increased productivity and profits, and were later kept in place by the firm. Likewise, the modern managerial practices introduced by Bloom, Eifert, Mahajan, McKenzie, and Roberts (2011) at large textile firms in India increased profits and were kept in place after the end of the experiment. In both cases, the owners attributed the failure to explore these options earlier as due to the high opportunity cost of their time. The same constraint seems to be binding both in very competitive environments, as experienced by the soft fruit firm we analyzed in Bandiera, Barankay, and Rasul (2007), and in settings where competition is very mild, as for the firms surveyed by Bloom, Eifert, Mahajan, McKenzie, and Roberts (2011).

This finding suggests promising new avenues for future research on the importance of time constraints in firms. The importance of time constraints at the top tiers of organizational hierarchies has been recognized in theory (Bolton and Dewatripont, 1994; Garicano, 2000), but evidence on whether and how managers allocate their time to maximize firm performance is scant. Bandiera, Guiso, Prat, and Sadun (2011) have developed a survey methodology to measure how chief executive officers spend their time, and show how the pattern of time allocation can be used to provide observational evidence on the internal and external constraints faced by firms. Combining this survey methodology with field experiments holds potential for making substantial progress in understanding these issues.

In turn, scarcity of managerial time can be symptomatic of two other problems. First, if the owner or chief executive officer must control all aspects of the business, the scope and size of the firm is necessarily limited. Delegation of authority and decision making is an essential ingredient for firm expansion, and yet we have a very limited empirical understanding of why some owners fail to delegate. Agency problems and the inability to motivate lower-tier managers intuitively seem important, as reported by the firm owners surveyed by Bloom, Eifert, Mahajan, McKenzie, and Roberts (2011), but more evidence is needed on how these can be tackled. Field experiments that vary the distribution of authority or the agency constraints could potentially provide this. Second, there is often a lack of managers with adequate human capital and talent, whether due to a market failure in education or a skewed distribution of talent. These constraints are not amenable to experimental variation, but by providing evidence on the internal constraints of firms, field experiments can help guide research on these topics, too.

Our discussion so far has been based on the assumption that firms maximize profits subject to constraints. But rather than focusing on constraints to optimizing behavior, a growing body of observational studies suggests the alternative view that firms might not maximize profits, either because they are led by managers who enjoy the “quiet life” (Bertrand and Mullainathan, 2003) or because they are owned by families whose objective function has a nonmonetary component deriving from direct control (Bertrand and Schoar, 2006). A burgeoning body of work uses field experiments to understand consumer preferences and optimizing behavior. It is worth exploring whether similar strategies could be adopted to shed light on departures from the assumption of profit maximization by firms.

Field experimenters have also begun to explore the behavior of not-for-profits. While there exists a vibrant literature using field experiments on the fund-raising activities of such organizations, many other issues remain unexplored. Theoretical contributions from Benabou and Tirole (2003) and Besley and Ghatak (2005) make clear that the provision of incentives for pro-social tasks raises different issues than for private tasks on at least two dimensions: First, to the extent that agents engaged in pro-social tasks are intrinsically motivated, financial rewards that might successfully elicit effort for private tasks could actually reduce effort if the rewards crowd out intrinsic motivation. Second, the type of incentive mechanism might affect the composition of the pool of agents who choose to participate in the activity. In particular, high-powered financial incentives might attract individuals who are motivated by financial returns instead of individuals who share the pro-social orientation of the organization, with undesirable consequences. Laboratory experiments show that the effect of financial rewards differs when the task has social value. But field evidence on these issues is scant.

Another question, with many practical implications, is how do firms and individuals at different layers of the hierarchy match endogenously based on their respective characteristics? The fact that many observed outcomes can be ascribed to endogenous matching is often a limitation of observational studies (Akerberg and Botticini, 2002). Field experiments on discrimination provide some evidence on how firms hire workers, but many questions remain open. We envisage field experiments that create exogenous variation in the parameters of the matching process, by, for instance, varying the information set available to employers and employees, or by reducing search costs through the introduction of electronic market places where employers and employees can meet. Advancing the methodology in these directions, perhaps in some cases intervening at the level of markets as a whole, would take field experiments into an exciting new realm.

■ *We have benefited from discussions with Nava Ashraf, David Autor, Stefano DellaVigna, Chad Jones, John List, Stephan Meier, Andrea Prat, David Reiley, Timothy Taylor, and seminar participants at the Econometric Society Meeting in Denver, 2011.*

References

Akerberg, Dan, and Maristella Botticini. 2002. "Endogenous Matching and the Empirical Determinants of Contract Form." *Journal of Political Economy*, 110(3): 564–91.

Aghion, Philip, and Jean Tirole. 1997. "Formal and Real Authority in Organizations." *Journal of Political Economy*, 105(1): 1–29.

Alonso, Ricardo, Wouter Dessein, and Niko

- Matouschek.** 2008. "When Does Coordination Require Centralization?" *American Economic Review*, 98(1): 145–79.
- Ashraf, Nava, Oriana Bandiera, and Kelsey Jack.** 2011. *No Margin, No Mission? A Field Experiment on Incentives for Pro-Social Tasks*. Unpublished paper, London School of Economics.
- Bandiera, Oriana, Iwan Barankay, and Imran Rasul.** 2005. "Social Preferences and the Response to Incentives: Evidence from Personnel Data." *Quarterly Journal of Economics*, 120(3): 917–62.
- Bandiera, Oriana, Iwan Barankay, and Imran Rasul.** 2007. "Incentives for Managers and Inequality among Workers: Evidence from a Firm Level Experiment." *Quarterly Journal of Economics*, 122(2): 729–74.
- Bandiera, Oriana, Iwan Barankay, and Imran Rasul.** 2009. "Social Connections and Incentives in the Workplace: Evidence from Personnel Data." *Econometrica*, 77(4): 1047–94.
- Bandiera, Oriana, Iwan Barankay, and Imran Rasul.** 2010. "Social Incentives in the Workplace." *Review of Economic Studies*, 77(2): 417–58.
- Bandiera, Oriana, Iwan Barankay, and Imran Rasul.** 2011. "Team Incentives: Evidence from a Field Experiment." <http://www.homepages.ucl.ac.uk/~uctpimr/research/teams.pdf>.
- Bandiera, Oriana, Luigi Guiso, Andrea Prat, and Raffaella Sadun.** 2011. "What Do CEOs Do?" CEPR Discussion Paper 8235.
- Barankay, Iwan.** 2010. "Rankings and Social Tournaments: Evidence from a Field Experiment." http://www.management.wharton.upenn.edu/~barankay/Documents/Rankings_and_Social_Tournaments.pdf.
- Benabou, Roland, and Jean Tirole.** 2003. "Intrinsic and Extrinsic Motivation." *Review of Economic Studies*, 70(3): 489–520.
- Bertrand, Marianne, and Sendhil Mullainathan.** 2003. "Enjoying the Quiet Life? Corporate Governance and Managerial Preferences." *Journal of Political Economy*, 111(5): 1043–75.
- Bertrand, Marianne, and Sendhil Mullainathan.** 2004. "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review*, 94(4): 991–1013.
- Bertrand, Marianne, and Antoinette Schoar.** 2006. "The Role of Family in Family Firms." *Journal of Economic Perspectives*, 20(2): 73–96.
- Besley, Timothy J., and Maitreesh Ghatak.** 2005. "Competition and Incentives with Motivated Agents." *American Economic Review*, 95(3): 616–36.
- Besley, Timothy J., and Maitreesh Ghatak.** 2008. "Status Incentives." *American Economic Review*, 98(2): 206–11.
- Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts.** 2011. "Does Management Matter? Evidence from India." NBER Working Paper 16658.
- Bloom, Nicholas, Raffaella Sadun, and John van Reenen.** 2010. "Does Product Market Competition Lead Firms to Decentralize?" *American Economic Review*, 100(2): 434–38.
- Bloom, Nicholas, and John Van Reenen.** 2010a. "Human Resource Management and Productivity." Forthcoming in the *Handbook of Labor Economics*, ed. O. Ashenfelter and D. Card. Elsevier.
- Bloom, Nicholas, and John Van Reenen.** 2010b. "Why Do Management Practices Differ across Firms and Countries?" *Journal of Economic Perspectives*, 24(1): 203–24.
- Bolton, Patrick, and Mathias Dewatripont.** 1994. "The Firm as a Communication Network." *Quarterly Journal of Economics*, 109(4): 809–39.
- Brown, Jennifer, and John Morgan.** 2009. "How Much is a Dollar Worth? Tipping versus Equilibrium Coexistence on Competing Online Auction Sites." *Journal of Political Economy*, 117(4): 668–700.
- Chiappori, Pierre-Andre, and Bernard Salanié.** 2003. "Testing Contract Theory: A Survey of Some Recent Work." Chap. 4 in *Advances in Economics and Econometrics Theory and Applications: Eighth World Congress Volume I*, ed. by M. Dewatripont, L. P. Hansen, and S. J. Turnovsky. Cambridge: Cambridge University Press.
- Cho, Sungjin, and John Rust.** 2010. "The Flat-Rental Puzzle." *Review of Economic Studies*, 77(2): 560–94.
- Coase, R. H.** 1937. "The Nature of the Firm." *Economica*, 4(16): 386–405.
- de Mel, Suresh, David McKenzie, and Christopher Woodruff.** 2008. "Returns to Capital in Microenterprises: Evidence from a Field Experiment." *Quarterly Journal of Economics*, 123(4): 1329–72.
- Drexler, Alejandro, Greg Fischer, and Antoinette Schoar.** 2011. "Keeping It Simple: Financial Literacy and Rules of Thumb." <http://personal.lse.ac.uk/fischer/Assets/KIS-DFS-Jan2011.pdf>.
- Garicano, Luis.** 2000. "Hierarchies and the Organization of Knowledge in Production." *Journal of Political Economy*, 108(5): 874–904.
- Harrison, Glenn W., and John A. List.** 2004. "Field Experiments." *Journal of Economic Literature*, 42(4): 1009–1055.
- Heckman, James J., and Peter Siegelman.** 1993. "The Urban Institute Audit Studies: Their Methods and Findings." In *Clear and Convincing Evidence: Measurement of Discrimination in America*, ed. M. Fix and R. Struyk, 187–258. Washington, D.C.: Urban Institute Press.
- Ibragimov, Rustam, and Ulrich K. Müller.** 2010. "t-Statistic Based Correlation and Heterogeneity

- Robust Inference." *Journal of Business and Economic Statistics*, 28(4): 453–68.
- Kandel, Eugene, and Edward Lazear.** 1992. "Peer Pressure and Partnerships." *Journal of Political Economy*, 100(4): 801–13.
- Karlan, Dean, and Martin Valdivia.** 2011. "Teaching Entrepreneurship: Impact of Business Training on Microfinance Clients and Institutions." *Review of Economics and Statistics*, 93(2): 510–27.
- Lazear, Edward P.** 1989. "Pay Equality and Industrial Politics." *Journal of Political Economy*, 97(3): 561–80.
- Lazear, Edward P.** 2000. "Performance Pay and Productivity." *American Economic Review*, 90(5): 1346–61.
- Lee, David S.** 2009. "Training, Wages and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *Review of Economic Studies*, 76(3): 1071–1102.
- Levitt, Steven D., and John A. List.** 2009. "Field Experiments in Economics: The Past, the Present, and the Future." *European Economic Review*, 53(1): 1–18.
- Levitt, Steven D. and John A. List.** 2011. "Was There Really a Hawthorne Effect at the Hawthorne Plant? An Analysis of the Original Illumination Experiments." *American Economic Journal: Applied Economics*, 3(1): 224–38.
- List, John A., and Imran Rasul.** 2011. "Field Experiments in Labor Economics." Chap. 2 in *Handbook of Labor Economics*, Vol. 4A, ed. O. Ashenfelter and D. Card. Elsevier.
- List, John A., Sally Sadoff, and Mathis Wagner.** 2010. "So You Want to Run an Experiment, Now What? Some Simple Rules of Thumb for Optimal Experimental Design." NBER Working Paper 15701.
- Lucas, Robert E.** 1978. "On the Size Distribution of Business Firms." *Bell Journal of Economics*, 9(2): 508–23.
- Mas, Alexandre, and Enrico Moretti.** 2009. "Peers at Work." *American Economic Review*, 99(1): 112–45.
- Mayo, Elton.** 1933. *The Human Problems of an Industrial Civilization*. New York: Macmillan.
- McKenzie, David.** 2011. "Beyond Baseline and Follow-up: The Case for More T in Experiments." Policy Research Working Paper 5639, World Bank.
- McKenzie, David, and Christopher M. Woodruff.** 2008. "Experimental Evidence on Returns to Capital and Access to Finance in Mexico." *World Bank Economic Review*, 22(3): 457–82.
- Moldovanu, Benny, Aner Sela, and Xianwen Shi.** 2007. "Contests for Status." *Journal of Political Economy*, 115(2): 338–63.
- Nagin, Daniel, James B. Rebitzer, Seth Sanders, and Lowell J. Taylor.** 2002. "Monitoring, Motivation, and Management: The Determinants of Opportunistic Behavior in a Field Experiment." *American Economic Review*, 92(4): 850–73.
- National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research.** 1979. "The Belmont Report" aka: "Ethical Principles and Guidelines for the Protection of Human Subjects of Research." April 18. <http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html>.
- Niederle, Muriel, and Lisa Vesterlund.** 2007. "Do Women Shy Away from Competition? Do Men Compete Too Much?" *Quarterly Journal of Economics*, 122(3): 1067–1101.
- Ostrovsky, Michael, and Michael Schwarz.** 2009. "Reserve Prices in Internet Advertising Auctions: A Field Experiment." <http://faculty-gsb.stanford.edu/ostrovsky/papers/rp.pdf>.
- Prendergast, Canice.** 1999. "The Provision of Incentives in Firms." *Journal of Economic Literature*, 37(1): 7–63.
- Reiley, David H.** 2006. "Field Experiments on the Effects of Reserve Prices in Auctions: More Magic on the Internet." *RAND Journal of Economics*, 37(1): 195–211.
- Riach, Peter A., and Judith Rich.** 2002. "Field Experiments of Discrimination in the Market Place." *Economic Journal*, 112(483): F480–F518.
- Rosen, Sherwin.** 1982. "Authority, Control, and the Distribution of Earnings." *Bell Journal of Economics*, 13(2): 311–23.
- Rotemberg, Julio J.** 1994. "Human Relations in the Workplace." *Journal of Political Economy*, 102(4): 684–717.
- Shearer, Bruce S.** 2004. "Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment." *Review of Economic Studies*, 71(2): 513–34.
- Shi, Lan.** 2010. "Incentive Effect of Piece-Rate Contracts: Evidence from Two Small Field Experiments." *B.E. Journal of Economic Analysis & Policy*, vol. 10, issue 1 (Topics), Article 61.
- Wu, Yanhui.** 2011. "Authority, Incentives and Performance: Theory and Evidence from a Chinese Newspaper." <http://personal.lse.ac.uk/WUY9/JMP.pdf>.