Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Market Design, Human Behavior, and Management

Yan Chen, Peter Cramton, John A. List, Axel Ockenfels

informs
http://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Articles in Advance, pp. 1–32
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Market Design, Human Behavior, and Management

Yan Chen,[a,b] Peter Cramton,[c] John A. List,[d] Axel Ockenfels[c]

[a] School of Information, University of Michigan, Ann Arbor, Michigan 48109; [b] Department of Economics, School of Economics and Management, Tsinghua University, Beijing 100084, China; [c] Department of Economics, University of Cologne, 50923 Köln, Germany; [d] Kenneth C. Griffin Department of Economics, University of Chicago, Chicago, Illinois 60637

Contact: yanchen@umich.edu, https://orcid.org/0000-0003-1287-4525 (YC); pcramton@gmail.com, https://orcid.org/0000-0001-8621-1092 (PC); jlist@uchicago.edu, https://orcid.org/0000-0002-1740-3848 (JAL); ockenfels@uni-koeln.de, https://orcid.org/0000-0003-1456-0191 (AO)

Received: September 20, 2019
Revised: March 4, 2020
Accepted: March 10, 2020
Published Online in Articles in Advance: September 9, 2020

https://doi.org/10.1287/mnsc.2020.3659

Abstract. We review past research and discuss future directions on how the vibrant research areas of market design and behavioral economics have influenced and will continue to impact the science and practice of management in both the private and public sectors. Using examples from various auction markets, reputation and feedback systems in online markets, matching markets in education, and labor markets, we demonstrate that combining market design theory, behavioral insights, and experimental methods can lead to fruitful implementation of superior market designs in practice.

History: Accepted by David Simchi-Levi, Special Issue of *Management Science:* 65th Anniversary.

Funding: This work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program [Grant 741409]; the National Science Foundation [Grants 1838994 and SES-1620319]; and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy [Grant EXC 2126/1-390838866].

Keywords: market design • human behavior • management

1. Introduction

Since the 1990s, economic research has played an increasingly important role in the practical organization and design of markets. The phrase *market design* includes "the design not only of marketplaces but also of other economic environments, institutions and allocation rules" (Roth 2015, p. 290). Prominent examples of market design include the auctions for spectrum, electricity, and other commodities; tradable permit systems for pollution abatement and other environmental regulations; online auctions; online reputation and feedback systems; financial markets; labor market clearinghouses; formal procedures for student assignment to public schools or colleges; centralized systems for the allocation of organs; and other related matching and trading processes. In many of these cases, theoretical, experimental and empirical research have complemented each other and influenced the design of market institutions.

In the process of bringing a theoretical idea or result to practice, the research strategy is often to observe the performance of the new market design in the context of the simple situations that can be created in a laboratory and assess its performance relative to what it was created to do and relative to the theory on which its creation rests. For this reason, laboratory experiments are often compared with a wind tunnel. For the rest of this section, we will briefly review several important papers published in *Management Science* related to market design and human behavior.

At the theoretical level, the most important tool for market design is game theory. In the first 20 years after von Neuman and Morgenstern published their seminal book, *Theory of Games and Economic Behavior*, game theory largely remained an academic pastime, primarily because of the technical difficulties of modeling games of incomplete information that underly almost all economic environments of interests (Morris 2019). Between 1967 and 1968, John Harsanyi published three path-breaking papers in *Management Science*, where he successfully argued that we can incorporate any incomplete information without loss of generality as the interim stage of some suitably constructed model of asymmetric information and extended Nash's concept of an equilibrium point to games of incomplete information (Harsanyi 1967, 1968a, b). One of the many important results from these papers was the concept of a *type* that summarizes the relevant characteristics of a particular player. These three papers provided economists with the much-needed tools for studying asymmetric information problems in strategic interactions (Gul 1997).

The first applied area of economics that embraced game theory was industrial organization, which generated many interesting insights on bargaining, contract design, pricing, and other practical problems that influenced the theory and practice of management. Game theory has since contributed considerably to virtually all applied theoretical research in economics and political science. Harsanyi's three *Management Science* papers are broadly considered the precursor to the game theory takeover of economic theory (Morris 2019). Primarily for his contributions in formalizing games of incomplete information, John Harsanyi, together with John Nash and Reinhardt Selten, received The Bank of Sweden Prize in Economic Science in Memory of Alfred Nobel in 1994 (Nobel Foundation 2019a).

In addition to theoretical foundations for market design, *Management Science* has also published a sequence of influential papers on human behavior. Here we highlight two such papers by researchers pivotal in the creation of the now vibrant field of behavioral economics. The first paper is from Kahneman and Lovallo (1993), who study choice under uncertainty by focusing on *isolation errors*, whereby people tend to treat risky prospects separately rather than together. In their first *prospect theory* paper, Kahneman and Tversky (1979) raised two central aspects of choice under uncertainty: the role of loss aversion and the probability weighting function. Isolation errors as the third component in risky choice is "something whose centrality to understanding risk attitudes researchers have only begun to fully appreciate" (Rabin 2003 page 169). In this paper, Kahneman and Lovallo not only presented experimental results demonstrating the prevalence of isolation errors but also applied it extensively in the context of managerial decision making to explain, for example, the pervasiveness of small-scale insurance policies, such as extended warranties on consumer products and the equity premium puzzle (Benartzi and Thaler 1995). "For having integrated insights from psychological research into economic science, especially concerning human judgment and decision-making under uncertainty," Daniel Kahneman shared the 2002 Nobel Prize in Economics (Nobel Foundation 2019b).

A second paper highlighted here is by Thaler and Johnson (1990), who investigate how risk-taking is affected by prior gains and losses. They present experimental data supporting the *house money effect*, whereby decision makers become more risk seeking in the presence of a prior gain, and *break-even effects*, whereby, in the presence of prior losses, outcomes that offer a chance to break even are especially attractive. Summarizing these empirical regularities, they propose an editing rule to describe how decision makers frame such problems. For having built "a bridge between the economic and psychological analyses of individual decision-making" and for his instrumental role "in creating the new and rapidly expanding field of behavioral economics," Richard Thaler received the Nobel Prize in Economics in 2017 (Nobel Foundation 2019c).

Finally, *Management Science* has published a sequence of papers on market design that combines theoretical insights with laboratory experiments to shed light on new market designs. Here we highlight Katok and Roth (2004), who investigate in the laboratory the performance of two auction formats for selling multiple homogenous objects: the ascending auctions used in eBay and the descending auctions best known for its use in the flower auctions in the Netherlands. The authors design three environments that include synergies and potentially subject bidders to the exposure problem and the free-riding problem. They find that the descending auctions perform well across environments, whereas the eBay ascending auction better avoids the free-riding problem. These findings have significant implications for market design for procurement and privatization. One year later, in 2005, *Management Science* published a special issue on Electronic Markets (Volume 51, Issue 3), which includes foundational auction design papers by various economists and computer scientists. Alvin Roth, together with Lloyd Shapley, received the 2012 Nobel Prize in Economics "for the theory of stable allocations and the practice of market design" (Nobel Foundation 2019d).

As demonstrated in these examples, *Management Science* has published foundational work in game theory, human behavior, and market design. Compared with mechanism design, which focuses on the use of game theory to understand how to efficiently design institutions, markets, and contracts respecting individual incentives, market design deals with a similar question but recognizes that theory can only go so far because many people are not (traditionally) rational or a necessary assumption of the theory means that critical things are left out. In the auction literature, the Vickrey-Clarke-Groves mechanism (Vickrey 1961, Clarke 1971, Groves 1973) is the output of the mechanism design approach, whereas ascending package bidding auctions are the output of the market design approach. Market design at its best takes the insights from game theory, behavioral economics, experiments, and field data to come up with practical institutional designs that have a real chance of improving existing institutions. Specifically, market design has a few distinguishing features compared with mechanism design. First, the objective of market design is to find institutions that work better. Second, market design emphasizes areas of inquiry where theory is relatively silent or underdeveloped. Last, market design should result in new (hopefully applied) mechanisms.

For the rest of the paper, we will survey several market design challenges and solutions, including strategic timing in auction and financial markets (Section 2), reputation and feedback system design in online markets (Section 3), matching market design in education (Section 4), and the design of labor markets (Section 5). Finally, Section 6 concludes.

2. Strategic Timing in Markets

Although economic theory simplifies markets and often does not worry about strategic timing, it is an important concern in market design. In matching markets, Roth and his coauthors analyze and develop mechanisms that address problems arising from incentives to act earlier than others (Roth 1990, 1991; Mongell and Roth 1991; Roth and Xing 1994; Roth and Peranson 1999; Kagel and Roth 2000; Roth 2008 provides a survey). Competition for people and positions in various job markets led to earlier and earlier dates of appointment, to the point that students were being hired before useful information about their performance was available and before the students themselves could develop informed career preferences. Roth designed and helped implement successful centralized matching algorithms to stabilize such markets (Roth 2002; Roth and Wilson 2019 provide an account of the history of market design and of recent developments).

Timing is also an important aspect of strategic behavior in auction markets. As we show in this section, a strategy called *sniping* (bidding as early or as late as possible to gain an advantage) is prevalent in many auction market environments, hampering the efficiency of trade. Sniping has probably been first observed in candle auctions, which were used starting about 1490 (Cassady 1967). In modern markets, market design solutions that can help mitigate sniping are often available. First, we show that sniping is widespread on consumer-to-consumer (C2C) online markets like eBay yet can be largely mitigated by changing the rule by which the auctions end. We then sketch how sniping arises in spectrum auctions and can be addressed by activity rules designed to promote better price discovery. Finally, we describe the race for speed in financial markets, why it arises, and how it can make traders worse off and create inefficiencies and market instabilities. Here, too, innovative market design solutions are available.

2.1. Online Auctions

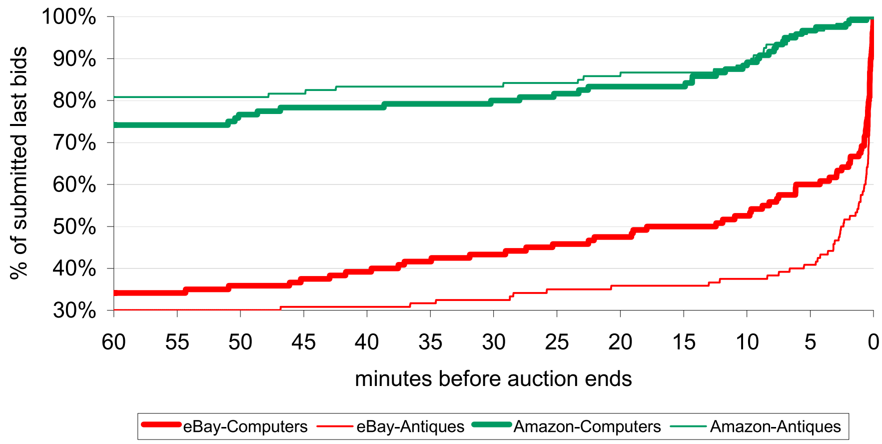
Many auctions, including online auctions for consumer goods, are often run in continuous time.[1] The simplest rule for ending such auctions is a fixed end time (a *hard close*), as used by eBay. A striking property of bidding on eBay is that a substantial fraction of bidders submits their bids in the closing seconds of an auction, which is called sniping, just before the hard close. Bidding is different on other platforms such as those formerly run by Amazon, which operated under otherwise similar rules. Amazon auctions were automatically extended if necessary past the scheduled end time until 10 minutes passed without a bid (a *soft close*).

Based on a study by Roth and Ockenfels (2002), Figure 1 shows the empirical cumulative probability distributions of the timing of the last bid in each auction for a sample of 480 eBay and Amazon auctions of antiques and computers with a total of 2,279 bidders. The timing of bids on Amazon is defined with respect to the initially scheduled deadline, which differs from the actual closing time if a bid comes in later than 10 minutes before the initial end time (only very few bids came in after the initially scheduled deadline, so we drop those observations for simplicity).

Figure 1 shows that there is significantly more late bidding on eBay than on Amazon. For instance, 40% of eBay computer auctions and 59% of eBay antiques auctions in the sample have last bids in the last five minutes compared with about 3% of both Amazon computer and Amazon antiques auctions that have last bids in the last five minutes before the initially scheduled deadline or later. The pattern repeats in the last minute and even in the last 10 seconds. This suggests that changes in the ending rules of auctions can strongly affect bidding behavior. Although the study of Roth and Ockenfels (2002) was one of the earliest on eBay, and the data were collected by hand, more recent studies of eBay referenced later use millions of auctions as data and mostly confirm the results.

Sniping on eBay is not easily explained by simple textbook auction analyses. The reason is that there is no time dimension in sealed-bid auctions, and dynamic auctions are typically modeled as clock auctions, where *price clocks*, instead of the bidding itself, determine the pace of the bidding. Moreover, eBay asks the bidders to submit maximum bids (called *proxy bids*). Because eBay's bidding agent will bid up to the maximum bid only when some other bidder has bid as high or higher, if the bidder has submitted the highest proxy bid, he wins at the *lowest possible price* of one increment above the next highest bid. Thus, similar to the second-price sealed-bid auction, at the end of the auction, a proxy bid wins only if it is the highest proxy bid, and the final price is the minimum increment above the second highest submitted proxy bid, regardless of the timing of the bid. This suggests that there is no reason to bid late. However, proxy bidding does not necessarily remove the incentives for sniping on eBay. Sniping can avoid bidding wars with incremental bidders, with like-minded

Figure 1. Cumulative Distributions over Time of eBay Auctions' Last Bids



Source. Reproduced from Roth and Ockenfels (2002).

late bidders, and with uninformed bidders who look to others' bids to determine the value of an item (see the series of papers by Roth and Ockenfels 2002 and Ockenfels and Roth 2006, 2013, that offer game-theoretic analyses for late and incremental bidding strategies and field evidence for strategic late bidding).

For example, sniping can be the best response to the late bidding strategies of like-minded bidders. In 2000, Hal Varian explained the underlying idea in a *New York Times* column titled "Online Users as Laboratory Rats" as follows: Suppose you are willing to pay up to $10 for a Pez dispenser, and there is only one other potential bidder who you believe also has a willingness to pay about $10. If both of you submit your value early, you will end up with a second highest submitted proxy bid of about $10, implying a price of about $10. Thus, regardless of whether you win or not, your earnings would be close to zero. Now consider a strategy that calls for a bidder to bid $10 at the very last minute and not to bid earlier, unless the other bidder bids earlier. If the other bidder bids earlier, the strategy calls for a bidder to respond by promptly bidding his true value. If both bidders follow this strategy and mutually delay their bids until the last minute, both bidders have positive expected profits, because there is a positive probability that one of the last-minute bids will not be successfully transmitted (Roth and Ockenfels 2002), in which case the winner only has to pay the (small) minimum bid. However, if a bidder deviates from this strategy and bids early, his expected earnings are (approximately) zero because of the early price war triggered by the early bid. Thus, with sniping, expected bidder profits will be higher and seller revenue lower than when everyone bids true values early. That is, sniping can be an equilibrium strategy even with private values and even if there is a risk that the snipe does not make it to eBay in time, before the auction closes.

When values are interdependent, there are additional strategic reasons to bid late in auctions, because the bids of others can then carry valuable information about the item's value that can provoke a bidder to increase his willingness to pay. This creates incentives to bid late, because by bidding late, less informed bidders can incorporate into their bids the information they have gathered from the earlier bids of others, and experts can avoid giving information to others through their own early bids (Bajari and Hortaçsu 2004, Ockenfels and Roth 2006, Hossain 2008).

Finally, last minute bidding can also be a best reply to incremental bidding. To see why, suppose you believe that your competitor starts with a bid well below his maximum willingness to pay and is then prepared to raise his proxy bid whenever he is outbid, as long as the price is below his willingness to pay. Last-minute bids can be a best response to this kind of incremental bidding because bidding near the deadline of the auction would not give the incremental bidder enough time to respond to being outbid. By bidding at the last moment, you might win the auction at the incremental bidder's initial, low bid, even when the incremental bidder's willingness to pay exceeds your willingness to pay. Nonstrategic reasons for incremental bidding include that bidders may not be aware of eBay's proxy system and thus behave as if they bid in an ascending (English) auction, *endowment effect* (Roth and Ockenfels 2002, Wolf et al. 2005, Cotton 2009), *auction fever* (Heyman et al. 2004), escalation of commitment and competitive arousal (Ku et al. 2005), uncertainty over one's own private valuation (Rasmusen 2006), or an unwillingness to reveal one's valuation (Rothkopf et al. 1990). Strategic reasons include shill bidding by confederates of the seller to push up the price beyond the second-highest maximum bid (Engelberg and Williams 2009) and a strategic response to the multiplicity of listings of similar objects (Anwar et al. 2006, Peters and Severinov 2006).

Amazon auctions are automatically extended if necessary past the scheduled end time until 10 minutes have passed without a bid. Although the risks of last-minute bidding remain, the strategic advantages of last-minute bidding are eliminated or severely attenuated in Amazon-style auctions, because no matter how late a bid was placed, other bidders will have time to respond. Thus, on Amazon, an attentive incremental bidder, for example, can respond whenever a bid is placed. As a result, the advantage that sniping confers in an auction with a fixed deadline is eliminated or greatly attenuated in an Amazon-style auction with an automatic extension (Ockenfels and Roth 2006, Malaga et al. 2010). Indeed, Figure 1 suggests that late bidding arises in large part from the rational response of the bidders to the strategic environment. Moreover, more experienced bidders on eBay bid later than less experienced bidders, whereas experience in Amazon has the opposite effect (Wilcox 2000, Ariely et al. 2005, Ockenfels and Roth 2006). In addition, because significantly more late bidding is found in antiques auctions than in computer auctions on eBay, but not on Amazon, behavior responds to the strategic incentives created by the possession of information in a way that interacts with the rules of the auction.

Laboratory experiments conducted by Ariely et al. (2005) replicate the major field findings in a controlled laboratory private-value setting in which the only difference between auctions is the ending rule. Moreover, the laboratory Amazon condition turns out to be more efficient and to yield higher revenues than the other conditions; the field evidence on efficiency and revenues from various auction platforms is, however, somewhat more mixed (Houser and Wooders 2005, Brown and Morgan 2009, Elfenbein and McManus 2010, Carpenter et al. 2011, Glover and Raviv 2012, Gray and Reiley 2013, Cao et al. 2019). Backus et al. (2015) find another harmful impact of sniping based on eBay field data: being sniped discourages new bidders from returning to bid again—they are between 4% and 18% less likely to return to the platform.

The next sections describe two other important examples for sniping in markets, examples in which traders are, unlike on eBay, most sophisticated and in which very different solutions to address sniping have been devised.

2.2. Spectrum Auctions

Spectrum auctions have been used by governments to assign and price spectrum for 25 years.[2] The development and implementation of innovative spectrum auction formats is among the greatest successes of market design. Over the years, many design issues have surfaced. Like on eBay (which was founded in 1995, around the same time when spectrum auctions started to become popular), one important challenge is to address incentives bidders may have to withhold expressing true demands until late in the auction and thereby undermine price discovery.

The workhorse for spectrum auctions since 1994 has been the simultaneous ascending auction, which is a simple generalization of the English auction to multiple items in which all items are auctioned simultaneously. Thus, unlike Sotheby's or Christie's auctions in which the items are auctioned in sequence, here all the items are auctioned at the same time: Each item has a price that is associated with it. Over a sequence of rounds, bidders are asked to raise the bid on any items that they find attractive, and the auctioneer identifies the provisional winner for each item at the end of every round. The process continues until nobody is willing to bid any higher, which is related to Amazon's soft close auction.[3] This process was originally proposed by Preston McAfee, Paul Milgrom, and Robert Wilson for the Federal Communications Commission (FCC) spectrum auctions.

Although these auctions end with a soft close, bidders may want to hold back, not pushing up prices on those objects they value most and concealing their private information until the end of an auction. One motivation for this strategy stems from an aggregate budget constraint. It may be easier to push a competitor aside late in the auction when the competitor has already committed its budget in other markets. A second motivation is a desire to better understand prices before committing to a specific portfolio of spectrum assets.

Sniping, however, slows the auction down and prevents price discovery.[4] Yet, good price discovery is essential in realizing the benefits of complex, dynamic auctions. One reason is that there is much uncertainty about what the objects being sold are worth. The bidders typically can only develop a crude valuation model. They need the benefit of some collective market insights, which can be revealed in a dynamic auction process to improve their bidding. If the price discovery process works well, the bidders gradually have their sights focused on the most relevant part of the price space. Focusing bidder decisions on what is relevant is probably the biggest source of benefit from the dynamic process (although this benefit is often ignored by economists, because economists typically assume that bidders fully understand their valuation models, when in practice bidders almost never have a completely specified valuation model). For such reasons price discovery is a public good and thus sniping, free-riding on others' efforts to find market prices, is a reasonable strategy if not prevented by auction design.

The standard solution in spectrum auction design is an *activity rule*. The activity rule requires a bidder to be active (that is to be the current high bidder or to submit new bids) on a predetermined quantity of spectrum licenses. If a bidder falls short of the required activity level, the quantity of licenses it is eligible to buy shrinks. Thus, bidders are prevented from holding back. The activity rule avoids late bidding and controls the pace of auctions by creating pressure on bidders to bid actively from the start. Milgrom and Wilson designed an activity rule that was applied to the U.S. spectrum auctions (McAfee and McMillan 1996, Milgrom 2004). Nearly all high-stake auctions, such as the FCC spectrum auctions, have an activity rule.

The exact design of the activity rule depends on the auction environment. More complex auctions require more complex activity rules. Too strong activity rules might force bidders to bid for less than their true demands, and too weak activity rules will inevitably lead to late bidding. For a single-object spectrum auction, a reasonable activity rule would require that no bidder can re-enter after exiting the auction. In an eBay-like auction, for instance, the activity rule would imply that all bidders, right at the start, submit their maximum willingness to pay as a proxy bid. No bidder could enter the auction once it started or re-enter once the bidder exited. (This, of course, would be incompatible with the flexibility needed on C2C auction platforms, but it is compatible with spectrum auctions where there are discrete rounds that follow a daily schedule.) For a multiunit auction of a single product, the activity rule would require that one cannot increase demand as price increases. For many related products, an aggregate quantity rule is needed, which requires that bidders are active on a particular fraction of current eligibility or the eligibility is reduced.[5] In more complex auctions, such as combinatorial clock auctions, state-of-the-art revealed preference rules can make sure that, as prices increase, bidders can only shift toward packages that become relatively cheaper (Ausubel et al. 2006, Ausubel and Baranov 2019).

What happens without an activity rule can be observed in spectrum auctions such as the Italy 4G auction, which did not have an activity rule. As a result, bidders held back demand, slowing the auction and limiting price discovery. Eventually, the auction lasted 470 rounds. That said, Germany's recent 5G auction, in 2019, lasted 497 rounds and thus set a new world record with respect to number of rounds in a simultaneous ascending auction. Here, the flaw was not the activity rule but the fact that it would take many rounds to get a one-increment increase in price, because Germany used the simultaneous ascending auction with bidding on individual lots rather than a clock auction, which has prices increase by a bid increment in each round for any product with excess demand (see Cramton and Ockenfels 2017 for an analysis of the German spectrum auction design). Measures to address sniping cannot be analyzed in isolation but must be closely connected to other details of the rules, such as pricing rules and increment rules to be fully effective.

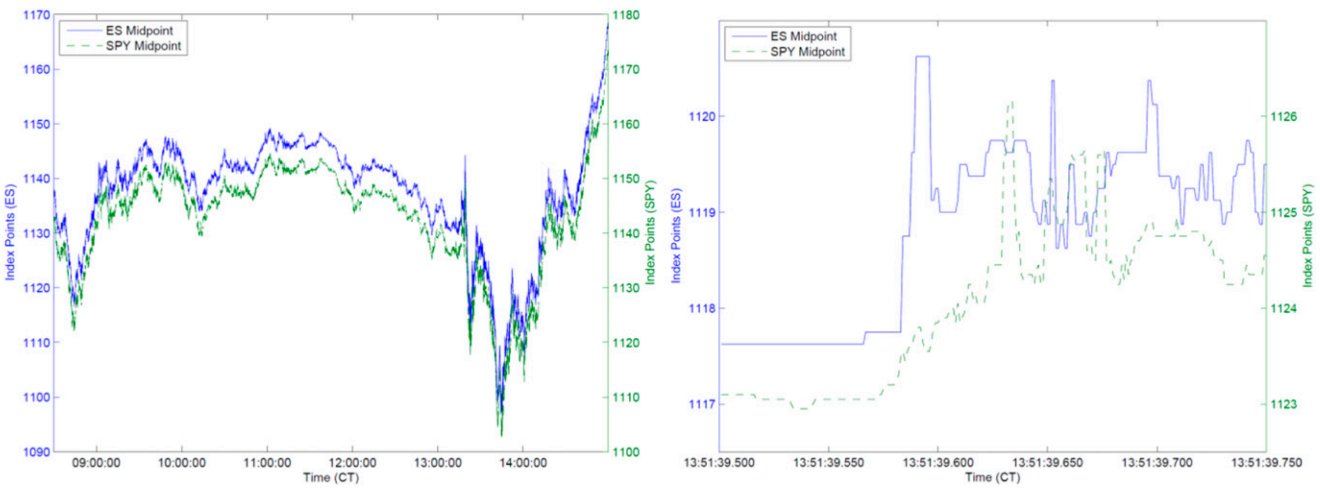
2.3. Financial Markets

Markets for financial securities are another important example where market design has a profound impact on the incentives for sniping and speed in markets. Unlike in spectrum auctions, the problem is not that bids tend to be held back but rather a never-ending arms race for ever faster trading. Because trading is continuous and equally attractive orders are processed in the order they arrive, speed is crucial in this format. This limits the performance of these markets (Budish et al. 2019). As before, the problem can be viewed with the lens of market design. This reveals a solution as presented in Budish et al. (2015), which we describe later.

The root of the problem is a fundamental flaw in today's markets: continuous-time trading. Continuous-time trading means that it is possible to buy or sell securities at any instant, where instant is measured in billionths of seconds—the speed of today's computers. Thus, the solution is for trading to occur in discrete time. Instead of trading at any instant, trading occurs, say, once per second. Orders arriving in the same second are batched together without any priority for orders that arrive a bit earlier, and all trades occur at the same price where supply and demand cross. The key is that the trading interval should be short as perceived by humans but long for a computer.

What exactly is wrong with continuous trading? Is trading as fast as possible not just good for price discovery and healthy competition, as probably suggested by our discussion of the need for activity rules in spectrum auctions? The answer boils down to a combination of two market failures. The first market failure is that in times of algorithmic trading, continuous markets do not and cannot work as they should in continuous time. Equivalent securities with prices that move in lockstep at human time intervals have moments of significant divergence at high frequencies. This creates what economists call technical arbitrage opportunities: the kinds of opportunities that are not supposed to exist if the market is working properly. For example, the price of the S&P 500 futures contract in Chicago (ES) and the S&P 500 EFT in New York (SPY) should move in perfect lockstep, and to the human eye, they do (Figure 2, left). However, when we zoom in to high frequency, there are hundreds of opportunities a day to make nearly riskless money—buy low in New York and sell high in Chicago or vice versa (Figure 2, right). This adds up to about $75 million

Figure 2. S&P 500 Index in Chicago (ES) and New York (SPY) Minute by Minute (Left) and Millisecond by Millisecond (Right)



Source. Reproduced from Budish et al. (2015).
Note. The securities move in lockstep on human time scale but are uncorrelated on a millisecond time scale.

a year for high-frequency traders, and this is just one pair of securities. There are hundreds of other pairs just like it, and, in our fragmented U.S. equities markets, trades are even simpler: if a stock jumps up on NASDAQ, buy it low on New York Stock Exchange (NYSE).

The second market failure is that these technical arbitrage opportunities, which are a prize to whichever trader snaps them up the fastest, create a never-ending arms race for speed. This fight for the prize is why there are investments like the $300 million high-speed cable between New York and Chicago and why that cable is already obsolete. This is why there are armies of physics and computer science PhDs devoted to shaving millionths or billionths of seconds off trading times. This is also why there are exchanges renting colocation services and high-speed data feeds; that is their way of getting a piece of the prize. Here is a simple way to think about it: continuous-time trading creates a $10 billion prize, and then high-frequency traders, exchanges, and broker dealers all scramble to get their piece.

Ultimately the prize comes out of the pockets of investors. The reason is that the technical arbitrage opportunities harm liquidity; it is harder to provide quotes to investors if one is constantly worried that prices will change and one's stale quotes will get picked off before one can revise them. Therefore, markets are less liquid than they should be. For institutional investors, this means that trading large blocks of stock is costlier.

Discrete time directly addresses both market failures. With discrete time, one cannot make money from exploiting pricing discrepancies that many traders see at the same time, just by acting a billionth of a second faster. This stops the arms race for speed. Unhealthy competition for speed is transformed into productive price competition. Trades occur at the right price, the consensus of the market, rather than at stale quotes. High-frequency traders still will be able to make money but only if they take actual risk, provide liquidity, or are smarter than the rest of the market (they know something that the rest of the market does not). One no longer can make money just from being the fastest to respond to some commonly observed event.

Discrete time also makes computational sense. Continuous trading implicitly assumes that computers and communications are infinitely fast. Computers and communications are fast but not infinitely so. Discrete time respects these limits. Tiny speed discrepancies between the direct feeds and public feeds of exchange data are critical with continuous time. This issue goes away with discrete time.

Continuous time breeds constant change and heightened complexity, making markets vulnerable to instability. Discrete time simplifies markets and allows both traders and exchanges to focus on improvements that make trading smarter and safer.

Which market design works in the financial sector to address sniping is the topic of many current discussions. Discrete time has seen limited implementation, and alternative design solutions have been proposed. For example, in 2016, the U.S. Securities and Exchange Commission approved the Investors Exchange (IEX) to operate as a public securities exchange. A primary goal of the IEX market, which was founded in 2012 to provide an alternative trading system with delayed messaging (Aldrich and Friedman 2018), is to reduce potential advantages of high-frequency trading firms. Another alternative is randomization of order priority as used by Electronic Broking Services (EBS), the largest currency exchange in the world. Asymmetric speed bumps (delaying sniping orders but not order cancelations) are now common. Other innovative methods such as flow trading are also being studied (Kyle and Lee 2017, Budish et al. 2020).

2.4. Future Directions: Economic and Algorithmic Design and the Pace of Price Discovery

There are many opportunities and important challenges in auction and market design (for surveys, see Klemperer 2004; Milgrom 2004, 2019; Cramton et al. 2006; Greiner et al. 2012; Bichler and Goeree 2017). Many of those opportunities and challenges follow from the fact that most digital platforms allow market engineers much control over the design, implementation, and operation of markets regarding pricing, demand and supply expression, information feedback, timing of transaction, and many other market features. Moreover, economic and algorithmic design is increasingly asked to address social concerns that go beyond economic efficiency, such as privacy and fairness. Exciting work at the interface of economics and computer science attests to these developments (examples include Bichler et al. 2010, Milgrom 2017, Kearns and Roth 2019, Parkes and Seuken 2020; see also the next sections). As a result, market and algorithmic design is shaping virtually all facets of economic and social interaction in many areas: online marketplaces, financial exchanges, the sharing economy, and platforms of social exchange.

In this section, we show that controlling the pace of price discovery is one of the pressing topics in this new era of market design. Interestingly, this was not anticipated by auction theory but rather inspired by practical challenges, low market performance, and failed design attempts. Analyses of spectrum, online, and financial markets demonstrate that sniping can often be explained by equilibrium predictions. Much of the late bidding in C2C online auctions such as eBay, on the other hand, is often best explained by a strategic response to naïve, incremental bidding, yet it can also arise at equilibrium in both private- and common-value auctions.[6] Indeed, the effect of the fixed deadline is likely as large as it is because it rewards late bidding both when other bidders are sophisticated and when they are not. Market design must sometimes consider not only the equilibrium behavior that we might expect experienced and sophisticated players eventually to exhibit but also how the design affects behavior of inexperienced participants, as well as the interaction between sophisticated and unsophisticated human players and algorithmic bidding agents.

However, unlike in spectrum and online auctions, which have experimented with various auction architectures both in the laboratory and the field,[7] there is not much conclusive and clean causal empirical evidence yet that reveals the relative performance of financial market institutions and that can guide market design for financial securities, despite the fact that policy makers worldwide are already taking actions intended to discourage high-frequency trading. Zhang and Riordan (2011), Brogaard et al. (2014), Menkveld and Zoican (2014), Benos and Sagade (2016), and Benos et al. (2017), among others, provide evidence for the costs of aggressive sniping. However, this literature comes from minor variants of the standard financial market design and thus offers no direct evidence about the costs and benefits of other platforms, engineered to eliminate the dilemma, as in Budish et al. (2015) and Aquilina et al. (2020). Moreover, although there are three decades of studying financial markets in the laboratory (for surveys on experimental research in financial markets, see Friedman and Rust 1993, Friedman 2010, and Noussair and Tucker 2013), aside from particular episodes such as the *Flash Crash* (Aldrich et al. 2016), little is known about the impact of sniping in times of financial stress as opposed to normal times (but see Jagannathan 2019 for a step in this direction). However, Aldrich and López Vargas (2019) recently conducted a laboratory market design study on high-frequency trading that suggests that, relative to the continuous double auction, the frequent batch auction exhibits less predatory trading behavior, lower investments in low-latency communication technology, lower transaction costs, and lower volatility in market spreads and liquidity. More studies on how financial market design affects sniping, market stability, and market resiliency are necessary.

Also, many other markets, as they move to real-time interaction, already see or will likely see similar problems and thus require new clever market design solutions. As an example, think about electricity market design, where we are just starting to observe similar issues. One of the reasons is the increasing share of intermittent renewables, which puts enormous stress on the system and increases the risk of outages, so that both improved investment incentives for reserve generation capacity (Cramton and Ockenfels 2012, Cramton et al. 2013) and more liquid real-time trading is needed. However, because the trend toward algorithmic trading in continuous electricity markets will also lead to a wasteful race for speed, this is posing serious threats to the efficiency and reliability of these markets (Neuhoff et al. 2016). Moreover, compared with financial markets, things tend to be more complicated in electricity markets because of complementarities in electricity production (Wilson 2002 and Cramton 2017). For instance, the race for speed in electricity trading hampers efficient pricing of transmission, which is often done on a first-come-first-serve basis in intraday trading. Also, a race for speed complicates the formulation and consideration of multidimensional bids, which consider the nonconvex cost structure of electricity production.

Another interesting example for the relevance of timing in markets is auction design for continuous sponsored search in the Internet, where other undesired bidding timing phenomenon have been observed, such as bidding cycles with automated bidding agents and various attempts to address those (Edelman and Ostrovsky 2007; Edelman et al. 2007; Varian 2007, 2009; Athey and Ellison 2011; Levin 2013 provide a survey). Clearly, taming sniping and improving price formation will remain a critical aspect of market performance in modern market environments.

Technology gives market designers an unprecedented ability to design and operate markets to better achieve objectives. One might expect rapid marketplace innovation as a result. However, progress is often slowed from the inertia of the status quo. Research is needed that improves our understanding of why innovation is difficult and how barriers of innovation may be overcome (see Budish et al. 2019 for a study of these challenges in financial markets). Too often market inefficiencies stem not from a lack of knowledge on how to fix the problem but from barriers to adopting the needed innovation.

3. Reputation and Feedback System Design in Online Markets

The astonishing success of online market platforms such as eBay, Amazon, Uber, and Airbnb can be attributed to the ease in which one side of the market can find a match on the other market side, as well as to the fact that they provide reliable information about the trustworthiness of the trading partner. All markets require some minimum amount of trust, yet this is a particular challenge for online markets and sharing platforms, where trades are typically with strangers, geographically dispersed, and executed sequentially. To incentivize trustworthiness, most online platforms use a reputation-based feedback system, enabling traders to publicly post information about past transaction partners. These systems have been, and are being, engineered based on conceptual insights from game theory and behavioral sciences and with the help of laboratory and field studies (surveys include Dellarocas 2003, Bar-Isaac and Tadelis 2008, Greiner and Ockenfels 2009, Bolton and Ockenfels 2012, Ockenfels and Resnick 2012, Tadelis 2016, Gutt et al. 2019).

3.1. A Case Study in Engineering Trust

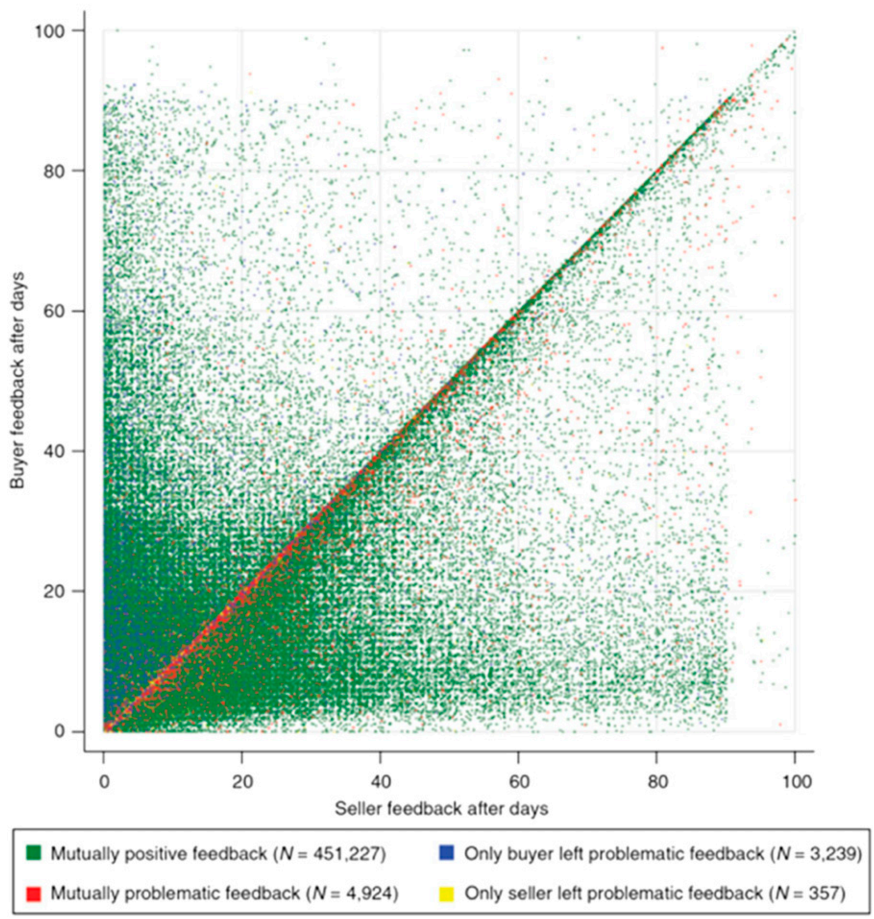
One major challenge of all feedback-based reputation systems is to get people to cooperate with the platform and leave feedback about their transaction partner. Feedback information is largely a public good, helping other traders to manage the risks involved in trusting an unknown transaction partner, so economists would tend to predict low participation rates. However, in the field data by Bolton et al. (2013), about 70% of the eBay traders, sellers, and buyers alike leave feedback (a number consistent with other research). It turns out that the key driver of provision of feedback, as well as the source of various distortions in feedback information identified in the literature, is *reciprocity*. More specifically, much of the feedback patterns we see can be organized by connecting them to two of the most fundamental research findings on the patterns of human cooperation in the last decades: altruistic punishment promotes cooperation, and counter-punishment hampers cooperation (Ostrom et al. 1992; Fehr and Gächter 2000a, 2002; Nikiforakis 2008; Mussweiler and Ockenfels 2013; Balafoutas et al. 2014). A natural way to (altruistically) punish a trader on an Internet platform who is not behaving according to what is perceived to be the social or trading norm is to leave negative feedback. This way, a propensity to altruistically punish norm-violators creates an incentive to be trustworthy. However, punishments can often be counter-punished, which is known to reduce the effectiveness of punishment to promote cooperation. Indeed, by retaliating a negative feedback with a negative one, counter-punishment may spoil the reputation of the altruistic punisher, which in turn may deter altruistic punishment in the first place. As a result, the potential of counter-punishment can hamper the effectiveness of reputation mechanisms and thus the performance of markets.

To illustrate the close analogy between (counter-) punishment in the behavioral science literature and giving feedback in the Internet, look at Figure 3, which is taken from Bolton et al. (2013). It shows the *timing* of feedback given on eBay by the buyer and the seller in hundreds of thousands of transactions. Most transactions either end with mutually positive (green dots) or with mutually negative feedback (red dots). Transactions with mutually positive feedback are all over the place (although a closer look at the data in Bolton et al. reveals that there is lots of reciprocity: many traders give kind feedback in reciprocal response to kind feedback). Transactions with mutual negative feedback, on the other hand, are highly clustered just below the diagonal. This means that many sellers, who are punished with a negative feedback from their buyers, respond *immediately* by counter-punishing with negative feedback. Clearly, feedback giving is not independent. The tightness and sequence in timing rather strongly suggest that sellers reciprocate positive feedback and retaliate negative feedback. Seller retaliation also explains why more than 70% of cases in which the buyer gives problematic feedback and the seller gives positive feedback (blue dots in Figure 3) involve the buyer giving second: the buyer going first would involve a high risk of retaliation. Observations in which only the seller gives problematic feedback (yellow dots) are rare and have their mass below the 45° line.

There are benefits and costs of reciprocity in feedback giving. A benefit of reciprocal positive feedback, for both the individual traders involved and the larger system, is that it helps getting mutually beneficial trades recorded. However, in the form of seller retaliation, reciprocal feedback imposes costs both on the buyers retaliated against and potentially on the larger system (because traders might not be willing to leave negative feedback out of fear that it will be retaliated). This would bias feedback information to be overly positive and therefore less informative in identifying problem sellers. Indeed, on eBay, almost all feedback is positive. Using internal eBay data, Nosko and Tadelis (2015) find that traders' positive feedback percentage is 99.3%, with a median of 100%. The concern is also supported by Dellarocas and Wood (2008), who examine the information hidden in the cases where feedback is *not* given. They estimate, under some auxiliary assumptions, that buyers are at least mildly dissatisfied in about 21% of all eBay transactions, which is far higher than the levels suggested by the reported feedback. They argue that many buyers do not submit feedback at all because of the potential risk of retaliation. Controlled laboratory evidence in Bolton et al. (2013) supports the notion that counter-punishment in feedback giving reduces the effectiveness of reputation building and market performance.

However, online reputation systems can be designed to address flaws in the system. Bolton et al. (2013) demonstrate that reciprocity can be guided by changing the way feedback information flows through the market system, leading to more accurate reputation information, more trust, and more efficient trade. Specifically, their data show that, compared with the simple two-sided feedback system traditionally implemented by eBay, where buyers leave feedback on sellers and vice versa, both *blind* and *one-sided* feedback significantly reduce the scope for retaliation, which in turn increase the informativeness of the feedback presented to buyers. The result is in line with game theory, behavioral science, laboratory and field research on social behavior and reputation building (such as the

Figure 3. Reciprocity in Feedback Giving



Source. Reproduced from Bolton et al. (2013).

line of research by Bolton et al. 2004, 2005; Bolton and Ockenfels 2009, 2014), and with field data collected across various market platforms. Indeed, the idea of making altruistic punishment easy but counter-punishment difficult explains important features of today's running reputational feedback systems. For instance, eBay supplemented their old two-sided feedback system with a one-sided system (called *detailed seller rating*). Based on research by Bolton et al. (2013), the one-sidedness was designed so that feedback cannot be retaliated by sellers. Airbnb, also inspired by the line of behavioral research described previously and their own experimental findings, created a blind feedback system, where transaction partners cannot see the others' feedback until they left their own. This also makes it impossible to retaliate negative feedback (although a recent study finds the effect to be small on Airbnb; Fradkin et al. 2019). Uber, on the other hand, makes it hard for passengers to identify a specific feedback giver, which is another way of making it difficult to retaliate negative feedback. Finally, eBay changed its systems again in 2008 so that sellers can only leave positive feedback, which was meant to eliminate the scope for counter-punishment.

3.2. Future Directions: Incentivizing, Filtering, and De-Biasing Human Judgment

There are important gaps in our knowledge, and more experimenting is needed to further improve trustworthiness and cooperation in online markets. For instance, because eBay's 2008 feedback system removes counter-punishment by sellers, *buyers* welcomed the new design (Klein et al. 2016). However, there are several indications that many *sellers* are unhappy with the new system. The reason is that by removing the option to counter-punish, the new system also removes the option to punish buyers and thus mitigates buyers' incentives to cooperate. To the extent that there is scope for moral hazard on the buyer side, this creates an imbalance of punishment (and thus bargaining) power between buyers and sellers. Thus, the overall effect of removing the sellers' punishment option on cooperation and market performance is probably more ambiguous than the study by Klein et al. (2016) suggests. From a broader perspective, the question how the rules affecting the scope for punishment and counter-punishment in interactions with *two-sided moral hazard* should be shaped largely remains an open one.

More recent attempts to incentivize, filter, and de-bias human judgment involve financial compensation for feedback information (see Li 2010, Li and Xiao 2014, and Li et al. 2016 for case studies on Alibaba, Cabral and Li 2015 for field experiments on eBay, and Burtch et al. 2018), plans to rely more on big behavioral data and artificial intelligence to better predict future behavior (Masterov et al. 2015, Luca and Zervas 2016, Milgrom and Tadelis 2018), and blockchain technology to better verify and audit transaction attributes (Catalini and Gans 2019).

Another pressing question in reputation design is whether and how traders can modify already submitted feedback information. One example is whether traders should be allowed to remove an initially given negative feedback if the dispute could later be resolved. Many platforms, including eBay.com, etsy.com, discogs.com, ricardo.ch, tradingpost.com.au, trademe.co.nz, mercadolibre.com, and listia.com, have or had a system that withdraws negative feedback from both traders' reputation profiles, if and only if both traders agree. Among other things, this option is thought to incentivize conflict resolution. However, Bolton et al. (2018) have shown both theoretically and empirically that this system is flawed in that it creates incentives to distrust, escalating conflict instead of resolving it. The reason is that the system allows traders to use counter-punishment to protect untrustworthy behavior: If I counter-punish a negative feedback that I received, my opponent will more likely agree to remove that negative feedback (because otherwise his reputation will also be spoiled). However, there is a lack of research that can guide the design of rules to integrate effective dispute resolution and informative reputation building systems (but see Ockenfels and Resnick 2012 and Bolton et al. 2020b).

There is also a lack of knowledge regarding feedback giving, content, and use in credence good markets, such as markets for medical, financial, or technical repair services. One major difference to the kind of online markets we have discussed this far is that consumers are often persistently unable to identify the quality of service that fits their needs best. This poses new challenges to designing effective and behaviorally robust mechanisms that promote trust and trustworthiness in these markets (Dulleck et al. 2011; Balafoutas et al. 2013, 2017; Kerschbamer et al. 2016, 2019).

We finally emphasize that research on *engineering trust* in online markets has been inspired by practical design problems. Indeed, standard reputation theory hardly anticipated the kinds of problems that online markets face when implementing reputation systems. Theory often assumes that reputation information is perfectly accurate and complete. Under these conditions, we can expect to see perfect reputation building and perfect trust among market actors (Wilson 1985, Bolton et al. 2011), and there is no scope for engineering literature that guides attempts to effectively promote the provision of informative feedback in practice. On the other hand, behavioral research and experimental studies turned out to be useful in organizing the relevant patterns observed in the field.

A desirable next step is to learn from such observations and develop new analytical models of the relevant institutional details and behavioral complexities in the field. For instance, although there has been much progress in modeling social behavior in the last two decades, including models of fairness and reciprocity (see Cooper and Kagel 2016 for an overview), as well as theoretical mechanism design implications of social preferences (Bierbrauer et al. 2017), no social preferences model captures the relevant punishment and counter-punishment patterns within an equilibrium framework (Engel 2014 and Dhami 2016 survey the literature). There is also comparatively little research on the psychological and social determinants of the production of reputation information, connecting the fundamental behavioral science literature on punishment and the practical market design literature on feedback giving. Interesting variables include the role of comparison processes for feedback giving and punishment (Chen et al. 2010, 2019; Mussweiler and Ockenfels 2013), social identity and discrimination (Chen and Li 2009, Cui et al. 2019, Kim et al. 2019, Bolton et al. 2020), and uncertainty (Ambrus and Greiner 2012, Bolton et al. 2019).

4. Matching Markets in Education

Although auction markets use prices to coordinate demand and supply, most of the centralized matching markets take agents' reported preferences as inputs and feed them into various matching algorithms to determine who gets what. Matching theory has been applied to many important design and management problems in both the private and public sectors, such as school choice, college admissions, course allocation, and entry-level labor markets. The practical design application of matching theory starts with the redesign of the National Residence Matching Program (Roth and Peranson 1997) and has since evolved into a research program that addresses practical market design problems using game theory, laboratory and field experiments, and computation methods (Roth 2002).

In what follows, we provide three examples of how a combination of economic theory and laboratory experiments informs the implementation of better education policies and management practices.

4.1. Redesign School Choice Mechanisms

School choice has been a widely debated education policy across the world, affecting the education experiences and labor market outcomes for millions of students each year.[8] The past two decades have witnessed major innovations in this domain. For example, shortly after the paper by Abdulkadiroğlu and Sönmez (2003) was published, New York City high schools replaced their allocation mechanism with a capped version of the student-proposing deferred acceptance (DA) mechanism (Gale and Shapley 1962), a less manipulable and more stable mechanism advocated by matching theorists involved in the design process (Abdulkadiroğlu et al. 2005a). In 2005, the Boston Public School Committee voted to replace its Boston immediate acceptance school choice mechanism (IA) with DA (Abdulkadiroğlu et al. 2005b) after IA was shown to be vulnerable to strategic manipulation both theoretically (Abdulkadiroğlu and Sönmez 2003, Ergin and Sönmez 2006) and experimentally (Chen and Sönmez 2006). In this case, experimental data helped make the case for DA in Boston's decision to switch from IA in 2005 (Abdulkadiroğlu et al. 2005b).

Within school choice research, matching mechanisms that have received significant scholarly attention include the Gale-Shapley Deferred Acceptance mechanism (Gale and Shapley 1962), the Boston Immediate Acceptance mechanism (Abdulkadiroğlu and Sönmez 2003), the Top Trading Cycles (TTC) mechanism (Abdulkadiroğlu and Sönmez 2003), and variants of the serial dictatorship mechanism (Pathak and Sönmez 2013). Indeed, the question of which mechanism best meets the goals of a school choice plan has been at the center of intensive research and ongoing policy discussions (Abdulkadiroğlu and Sönmez 2003, Ergin and Sönmez 2006, Abdulkadiroğlu et al. 2011).

We first briefly introduce the school choice mechanisms, summarize their theoretical properties, and describe performance in the laboratory and field when applicable. We then discuss the major school choice reforms around the world concerning the abandonment or adoption of some of these mechanisms.

Our first mechanism, IA, is the most common school choice mechanism observed in practice in China, the United Kingdom, and the United States. Its outcome can be calculated via an algorithm that puts a lot of emphasis on a student's reported top choice. In the first step of the algorithm, each school only considers students who have listed it as their top choice and sorts them in priority order. Each school admits those with the highest priority and rejects the rest. Those rejected enter the second step of the algorithm, and so on. The algorithm terminates when there are no students left to assign or no school seats remain. Importantly, the assignments in each step are final.

We will use a simple example to illustrate the incentive problems created by this algorithm. Consider a fictional student, Anna, who applies to elementary schools under the IA algorithm. There are three public elementary schools in her school district, Angell, Burns Park, and King. Anna lives in the Burns Park district, which gives her high priority at Burns Park and low priority elsewhere. Her top choice is Angell. Her second choice is Burns Park, and her third choice is King. If Anna ranks her preferences truthfully but does not get into

Angell (likely because she has lower priority there), her application will be sent to Burns Park. However, if all Burns Park seats are filled in the first round, Anna loses her priority advantage and is assigned to her last choice, King. In this case, we say Anna has *justified envy* as she is not assigned to Burns Park but she prefers Burns Park to her assignment, and she has a higher priority than some student who is assigned to Burns Park.

If she plays it safe and lists Burns Park as her top choice, she is guaranteed a seat at Burns Park, a better outcome than being assigned to King. Based on this feature, an important critique of IA is that it gives students strong incentives for gaming through misreported preferences. That is, a student who has high priority for a school under IA may lose her priority advantage for that school if she does not list it as her first choice. Consequently, IA forces students to make hard and potentially costly choices, which leads to a high-stakes game among participants with different levels of strategic sophistication. This has been observed in the laboratory among financially motivated subjects (Chen and Sönmez 2006) and in the field, using naturally occurring data from Boston (Pathak and Sönmez 2008). Recognition of these deficiencies of IA has lead Boston and many other cities in the United States to abandon IA and replace it with DA.

Outside of the United States, variants of IA have been used as a school choice mechanism in China, France, and the United Kingdom. In China, to equalize access to school resources across students of different socioeconomic backgrounds, the Chinese government abandoned the previous merit-based middle school admissions mechanism in 1998 and replaced it with an open enrollment school choice mechanism where parents rank schools and schools select students using IA (Lai et al. 2009). Since then, students applying for middle schools are prioritized based on their residence, whereas those applying for high schools are prioritized based on their municipal-wide examination scores. Using public middle school admissions data from Beijing Eastern City District, He (2014) investigates parents' behavior and finds that parents are overcautious in that they play safe strategies too often. Combining survey data, middle school choice data, and high school entrance examination test scores from Beijing, Lai et al. (2009) find that children of parents who made mistakes in middle school selection were admitted to lower-quality schools and achieved lower test scores on the high school entrance examination three years later. Despite these problems, IA continues to be used as the major school choice mechanism in China.

Our second mechanism is the student-proposing deferred acceptance mechanism (Gale and Shapley 1962), which has played a central role in the school choice reforms in Boston and New York City (Abdulkadiroğlu et al. 2005a, b), as well as in Finland, Ghana, Paris (Fack et al. 2019), Romania, Singapore, and Turkey. The outcome of this mechanism can be calculated via the *DA algorithm*. In the first step of the algorithm, each school also only considers students who have listed it as their top choice and sorts them in priority order. Each school put those with the highest priority *tentatively* on hold and rejects the rest. Those rejected apply to their second-choice school, which re-sorts those on hold from the previous round and the newcomers based on their priority, put those with the highest priority on hold and reject the rest, and so on. The algorithm terminates when each student is tentatively retained at some school. In DA, assignments made at each step are temporary until the last step. This feature contributes to the desirable properties of DA in terms of incentives and stability.

Consider our fictional student Anna's brother, Marco, who applies to elementary schools under the DA algorithm. He lists his choices in the true order of his preferences, which are the same as his sister's: Angell, Burns Park, and King. If Marco does not get into Angell either, his application will be sent to Burns Park. The algorithm then re-sorts everyone retained from the first round together with the newcomers based on their priorities. Because Marco does not lose his priority advantage, he is assigned to Burns Park. Therefore, truth telling does not hurt Marco and may sometime make him strictly better off.

To summarize, one advantage of DA is that it is strategy-proof (Dubins and Freedman 1981, Roth 1982). That is, when students can list as many choices as they want, DA allows them to safely rank schools in true order of preferences. They will not lose a place just because someone else applies earlier in the algorithm. A second advantage of DA is that it produces the stable matching that is most favorable to each student. In other words, when the algorithm finishes, there will not be any student and any school that are not matched with each other but that would *both* prefer to be. Although its outcome is not necessarily Pareto efficient, it is constrained efficient among the stable mechanisms.

In many laboratory experiments testing DA, researchers find that it remains the mechanism that achieves the highest proportion of stable allocations. Depending on the size of the match, the proportion of students revealing their preferences truthfully varies between 47% to more than 80% (Hakimov and Kübler 2019). In addition to Boston and New York City, variants of DA have been implemented in Amsterdam, Denver, Hungary, Paris, New Orleans, and Taiwan.

In the tradeoff between elimination of justified envy and Pareto efficiency, DA gives up Pareto efficiency. The TTC mechanism, on the other hand, gives up elimination of justified envy but is Pareto efficient.

In each round of the TTC algorithm, each student points to her favorite school among schools that remain, whereas each school points to the applicant who has the highest priority at that school among the remaining applicants. A cycle of students and schools pointing at each other is called a *top trading cycle*. Every student in a cycle is assigned to the school she is pointing to. These students and their assignments are removed from the allocation process. School capacity is updated. The algorithm terminates when each student is assigned a school seat or all school seats are assigned.

The TTC mechanism is not only Pareto efficient but also strategy-proof. In the laboratory, however, without prompting from the experimenters, sometime up to one third of the subjects manipulate their preferences (Chen and Sönmez 2006).

In theory, TTC has an efficiency advantage over IA and DA: The outcome of IA is Pareto efficient if participants reveal their preferences truthfully. Any efficiency loss in IA is a consequence of preference manipulation. DA, on the other hand, is strategy-proof, but elimination of justified envy and Pareto efficiency are not compatible. Because DA Pareto dominates any other mechanism that eliminates justified envy, any efficiency loss in DA is a consequence of this incompatibility.

In practice, the only public school district that implemented TTC as its school choice mechanism is the New Orleans Recovery School District (RSD). However, one year after its implementation, the RSD switched to DA, citing the difficulty to explain TTC to parents and the lack of stability as main reasons for the switch. Using data from New Orleans, Abdulkadiroğlu et al. (2017) find that the switch to DA had little impact on the overall aggregate rank distribution of choices received by applicants; however, no student is involved in a blocking pair as a result of the switch. Based on the revealed preferences of officials in both the Boston and New Orleans public schools, one cannot help but notice that they seem to put more weight on stability, guaranteed by DA, compared with efficiency, guaranteed by TTC. The lack of stability might create blocking pairs, which might lead to legal challenges to the school district.

Last, in the fall of 2009, without the involvement of market design researchers, Chicago Public Schools decided to replace its highly manipulable matching algorithm for exam schools, a variant of IA, with a less manipulable mechanism, a capped version of the serial dictatorship (Pathak and Sönmez 2013).

In sum, market design in the school choice domain is considered a success story, with many active research projects investigating school choice mechanisms around the world. In many cases, researchers are directly involved in the design of better matching algorithms, whereas in other cases, officials from public school districts abandon problematic matching algorithms in favor of less manipulable ones.

4.2. Redesign Centralized College Admissions Mechanisms

Like school choice, college admissions policies have been subject to debate and reform in many countries.[9] In particular, many countries use centralized college admissions through standardized tests, including China, Greece, Hungary, Russia, and Turkey. In what follows, we discuss the role of matching theory and experiments in the understanding the recent college admissions reforms in China.

In China, centralized matching processes via standardized tests assigning students to universities have been in place since 1952. The National College Entrance Examination forms the foundation for the current college admissions system. In recent years, roughly 10 million high school seniors compete for 6 million seats at various Chinese universities each year. The matching of students to universities has profound implications for the education and labor market outcomes for these students. Through its regional variations and its evolution over time, the Chinese system also provides matching theorists and experimentalists with a wealth of field observations to enrich our understanding of matching mechanisms.

In recent years, each province implements an independent matching process from one of the two classes of mechanisms: the sequential or the parallel mechanism. The sequential mechanism, strategically equivalent to the IA mechanism, had been the only mechanism used in Chinese student assignments both at the high school and college level until 2000 (Nie 2007). However, this mechanism is not strategy-proof: “a good score in the college entrance exam is worth less than a good strategy in the ranking of colleges” (Nie 2007, p. 23).

To alleviate the problem of high-scoring students not being accepted by any universities, the *parallel mechanism* (PA) was first implemented in Hunan Province in 2001. In the parallel mechanism, students select several parallel colleges within each choice-band. For example, a student’s first choice-band may contain a set of three colleges, A, B, and C, whereas her second choice-band may contain another set of three colleges, D, E, and F (in decreasing desirability within each band). Assignments for parallel colleges listed in the same band are considered temporary until all choices of that band have been considered. Thus, this mechanism lies between IA, where every choice is final, and DA, where every choice is temporary until all seats are filled.

By 2019, all 31 provinces and autonomous regions had abandoned the sequential in favor of various versions of the PA, which is widely perceived to improve allocation outcomes for students. These variants of PA are differentiated by the number of parallel colleges a student can list within a choice-band.

To investigate the theoretical properties of the PAs, Chen and Kesten (2017) formulate a parametric family of application-rejection mechanisms where each member characterized their parallel and periodic choice-band sizes that allow the application and rejection process to continue before assignments are made permanent. As the choice-band size increases, we go from IA to PA, and from those to DA. They show that members of this family become less manipulable in the sense of Pathak and Sönmez (2013) and more stable as the choice-band size increases. This implies that the Chinese provinces that have adopted a parallel mechanism have transitioned to a less manipulable and more stable assignment system.

Furthermore, Chen and Kesten (2017) show that a PA mechanism provides students with a certain sense of insurance by allowing them to list their equilibrium assignments under the IA mechanism as a safety option while listing more desirable options higher up in their preferences. This strategy leads to an outcome at least as good as that of the IA mechanism. Notably, such insurance does not come at any ex ante welfare cost in a stylized setting.

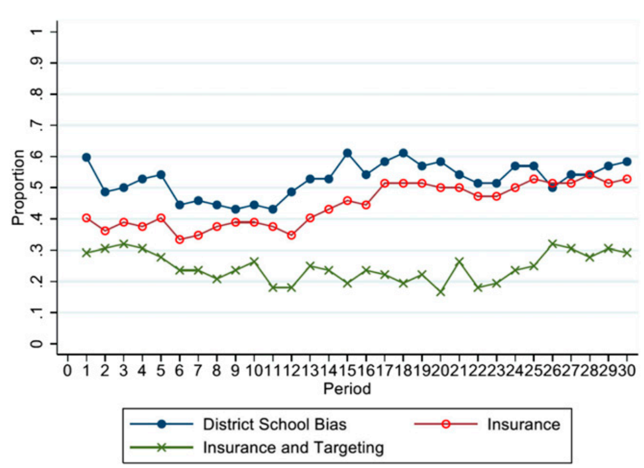
To investigate behavioral responses to these mechanisms and to search for behavioral regularities where theory is silent, Chen and Kesten (2019) evaluate the IA, PA, and DA mechanisms in the laboratory in two environments differentiated by their complexity. In the laboratory, participants are most likely to reveal their preferences truthfully under DA, followed by PA and then IA. Furthermore, although DA is significantly more stable than PA, which is more stable than IA, efficiency comparisons vary across environments. Whereas theory is silent about equilibrium selection, they find that stable Nash equilibrium outcomes are more likely to arise than unstable ones. Regardless of the metrics, the performance of PA is robustly sandwiched between IA and DA.

One type of strategy implied by proposition 5 in Chen and Kesten (2017) is an insurance strategy. In the six-school environment in Chen and Kesten (2019), students' district school position varies from the second to the fifth position. A student has a high priority in her district school and low priority elsewhere. To insure that a student gets a school at least as good as her district school, she can put her district school as her second choice and a more preferred school as her first choice, called an *insurance strategy*. Within this subset, if a student lists her most preferred school as her first choice and her district school as her second choice, we label it as the *insurance and targeting strategy*. Figure 4 presents the proportion of students adopting district school bias (DSB, ranking one's district school higher than its true position), insurance, and insurance and targeting strategies over time. By the last period, 58%, 53%, and 29% of the subjects adopt DSB, insurance, and insurance and targeting strategies, respectively. That is, 91.4% (respectively, 50%) of those who use DSB actually use the insurance (respectively, insurance and targeting) strategy, confirming both the popular perception and the theoretical characterization of the insurance property of PA. These results help explain the recent reforms in Chinese school choice and college admissions.

In practice, we observed changes within the parallel family. For example, Hunan Province pioneered the parallel mechanism in 2001, which allowed three parallel choices per choice-band. Later, it switched to a different parallel mechanism allowing five parallel choices per choice-band in 2010. Using admissions data from Hunan, Wei (2015) find that, by 2013, the new parallel mechanism with a choice-band size of five is significantly more stable than the old parallel mechanism with a choice-band size of three. In future studies, it would be desirable to pick more members to investigate the performance of different PA mechanisms.

Researchers have also examined other aspects of the Chinese college admissions mechanism, such as the timing of preference submission by students. Recent empirical and experimental studies such as Wu and Zhong (2014), Lien et al. (2016), and Jiang (2016) find that if students submit preferences before taking the examination, the measurement error in the examination can be corrected via IA, which leads to matchings that are stable with regard to students' aptitudes.

Figure 4. Proportion of Students Adopting District School Bias, Insurance, and Insurance and Targeting Strategies Under the Chinese Parallel Mechanism in Chen and Kesten (2019)



Source. Reproduced from Chen and Kesten (2019).

In sum, there has been a fair amount of progress made in understanding and testing various aspects of centralized college admissions mechanisms. Because of the large number of choices available for centralized college admissions, it remains a challenge to design optimal matching mechanisms that simultaneously provide information to guide the application process and reduce the students' cognitive load in preference reporting.

4.3. Improving Course Allocation Mechanisms

Allocating course seats to students is a daunting task each university face every semester. It is a technically difficult problem in market design, because it involves assigning each student a package of indivisible goods among many classes where some are substitutes and others are complements. The course allocation problem is closely tied to the combinatorial auction problem discussed in Section 2.2. A critical theoretical distinction is the presumption of quasilinear utility in the auction problem, which is typically not used in matching markets.

To achieve the goals of efficiency and equity, some business schools use preference-ranking mechanisms (revealed ordinal preferences), whereas others use variants of bidding systems (revealed cardinal preferences) in which students are given a fixed budget of tokens to bid on courses. In such bidding systems, bids serve the dual roles of inferring student preferences over courses and determining student priorities for courses. Similar market-like mechanisms with tokens have been proposed and tested to solve resource allocation problems for organizational computing (Ledyard 1993, (Olson and Porter 1994) and for scarce scientific and engineering resources (Takeuchi et al. 2010), as well as for prediction markets within firms (Arrow et al. 2008).

Sönmez and Ünver (2010) present a theoretical analysis of course bidding, where they show that these dual roles may easily conflict. That is, preferences inferred from the bids might differ significantly from students' true preferences. Furthermore, they propose the Gale-Shapley student-optimal stable mechanism (DA) that can be implemented by asking students for their preferences in addition to their bids over courses. The DA mechanism operates as follows in this context. In the first step, each student is tentatively placed in her top three choices from her preference list. If a course has more students than its capacity, students with the lowest bids for that course are dropped. A student rejected from a course is tentatively placed in her next choice course, and so on. The algorithm terminates when no student is dropped from a course, or all options on the students' preference list are exhausted. The tentative assignments become final.

To compare the new mechanism with the existing bidding system, Krishna and Ünver (2008) report a field study complemented by a laboratory experiment at the Ross School of Business at the University of Michigan, which uses a course-bidding system. In this system, each student is endowed with a fixed budget of bidding points, which they can allocate among courses they are interested in. Students are then sorted in decreasing order by the points they place in a course, which generates a priority list. A serial dictatorship mechanism is executed in priority order, subject to course quota and feasibility constraints. The field study involves a personalized email sent to each student within a few hours of the official closure of course bidding. The email contains a list of all courses on which the student had placed bids in descending order of bid points and asks students to rank the courses. Their counterfactual analysis using DA concludes that a potential transition to DA is likely to lead to significant efficiency improvement: among the 489 students who submitted a rank-ordered list, 101 of them unambiguously prefer DA, whereas 2 strictly prefer the status quo. The others are indifferent. In a complementary laboratory experiment using the induced value method, the authors again find an improvement in efficiency under DA. Depending on the metric used, truthful preference revelation under DA is between 67% and 83%. Despite the evidence that switching to DA would improve the satisfaction level of many students, the Ross School of Business at the University of Michigan continues to use the course-bidding system.

A second example of the course-bidding system is used by Harvard Business School. Budish and Cantillon (2012) document how it encouraged strategic behavior and often failed to produce efficient outcomes. To address these deficiencies, Budish (2011) proposes a new allocation mechanism that elicits from students their preferences over bundles of courses and uses these preferences to compute a price for each course that would form an approximate competitive equilibrium from equal income (A-CEEI). At these prices, each student receives her most preferred bundle of courses that she could afford. As the number of participants grows large, the amount of approximation and the incentives to misrepresent preferences would become small. The main advance of Budish (2011) compared with Sönmez and Ünver (2010) is to allow students to express preferences over bundles of courses rather than individual courses, thus capturing potential substitutability or complementarity among various courses.

To implement A-CEEI in practice, market designers need to deal with the issue of preference reporting over bundles of courses, which would be prohibitively large. A practical mechanism will necessarily use a simplified preference reporting language, which in turn raises the empirical question of how well the restricted preferences approximate true preferences. The first experimental investigation and subsequent implementation of A-CEEI took place in the Wharton Business School at the University of Pennsylvania. Before 2013, the Wharton Business School used a course-bidding system called the Wharton Auction.

In a paper in *Management Science*, Budish and Kessler (2019) report a novel laboratory experiment that compares the performance of A-CEEI with the existing Wharton Auction. There are several interesting features in the experiment design. For example, subjects are Wharton MBA students who have experience with the Wharton Auction and who would be the future users of A-CEEI if it were adopted. Instead of endowing these subjects with artificially induced values (Smith 1982), they bring their real preferences into the laboratory. Specifically, subjects report preferences over a subset of courses to be offered the following semester, making it a realistic task. An innovation in the preference reporting language is the use of binary comparisons, in the form of "Do you prefer Schedule A or Schedule B?," which is cognitively simple compared with ranking over all possible schedules. They find that A-CEEI outperformed the incumbent Wharton Auction on measures of efficiency and fairness.

The experimental results helped persuade the Wharton committee to adopt A-CEEI and guide its practical implementation. The new mechanism is implemented as Course Match, which has replaced the Wharton Auction since 2013. Survey data suggest that A-CEEI has increased student satisfaction with their assigned schedule (Budish et al. 2017).

4.4. Future Directions: Dynamic Mechanisms and Information Intervention

With the advent of the World Wide Web and the computerization of most matching markets, practitioners, with or without the help of market designers, sometimes introduce new matching mechanisms or new features to existing matching mechanisms that take advantage of the low cost of information gathering and dissemination over the Web. These new features, enabled by information technology, might shape the future of matching market design. Summarizing emergent research on matching markets around the world, we highlight two directions for future research.

The first direction is dynamic matching market design. Although most existing research in the school choice and college admissions domain analyzes the static version of the matching mechanisms, we notice dynamic variants of these mechanisms being implemented that provide information about others' behavior and allow students to revise their own applications on observing others' actions. Examples include school choice in Amsterdam (de Haan et al. 2016) and Wake County, North Carolina (Dur et al. 2018), where students (or parents) can revise their choice based on feedback on others' choices. In the context of the Japanese Residency Matching Program, applicants can check the number of students listing each hospital program as their first choice and revise their rank order list within a prespecified time window.

In the college admissions context, a more radical dynamic matching market starts to emerge in the form of the dynamically adjusted admission cutoffs at each college during a prespecified time frame in the college admissions process in Inner Mongolia (Gong and Liang 2019) and Brazil (Bó and Hakimov 2020). During this process, students can revise their applications continuously upon observing others' actions. Although research analyzing existing information provision mechanisms in matching demonstrates that it improves the performance of these mechanisms compared with their static counterparts, it remains an open question what the optimal dynamic matching mechanisms might be.

A second promising direction is related to information intervention before the matching stage. When choosing a school or college, students often have imperfect information about their own preferences regarding candidate schools, partly because it is difficult to assess the potential educational outcomes each school provides (Dustan et al. 2015). Unfortunately, acquiring this information can be costly if a student faces too many choices or must acquire information about several factors, such as academic performance, teacher quality, school facilities, extracurricular activities offered, and peer quality. Chen and He (2019, 2020) investigate how two popular mechanisms, IA and DA, incentivize students' information acquisition theoretically and in the laboratory. Although students' willingness to pay for information is significantly greater under IA than DA, as predicted by theory, most students' information acquisition behavior is far from optimal. Their counterfactual analyses show that providing information on both a student's own and others' preferences is welfare-enhancing. Furthermore, students who never invest in information acquisition benefit equally from information provision. In light of these results, the new dynamic matching mechanisms partially serve the function of information provision about others' preferences.

To reduce inequality in access to education resources, information intervention before students' entry into matching markets might be an effective policy, because the cost of information acquisition is particularly harmful to low-income students. Indeed, research has shown that, because of limited information, low-income high achievers in the United States tend not to apply to selective colleges, despite that generous financial aid makes these colleges more financially accessible than the colleges these students end up choosing (Hoxby and Avery 2013, Hoxby and Turner 2015). Information intervention can therefore substantially raise the number of applications from these students to selective colleges. In a natural field experiment conducted in the state of Michigan, Dynarski et al. (2018) contacted low-income, high-achieving public high school students, their parents, and principals with an encouragement to apply to the University of Michigan and a promise of four years of free tuition and fees on admission. They find that treated students were more than twice as likely to apply to and enroll at the University of Michigan, with no diversion from schools as (or more) selective as Michigan. Furthermore, this application tendency is not limited to college selection. Similar undermatching phenomena are observed among low-income families in public school choice plans (Hastings and Weinstein 2008). Future research on effective policies to reduce information frictions can lead to substantial welfare gain.

5. Labor Market

The labor market represents a rich assortment of opportunities for the scientist and manager to explore ways to improve existing conditions. Within economics, a myriad of research questions is addressed by labor economists today, yet the bulk of work revolves around three bins: (i) labor supply, (ii) labor demand, and (iii) the organization of labor markets and behavioral incentives therein. There is by now enough work in these three areas to fill at least a hundred factual tomes. In this section, we limit our attention to a subset of the third bin: exploring select recent work on the effects of market design and pecuniary incentives on worker behavior and the interplay between pecuniary and nonpecuniary incentives (the interested reader should see List and Rasul 2011 for a discussion of the other two bins). We then explore challenges moving forward and our vision of the next set of frontiers.

A core feature of economics is that incentives matter. The key is to understand what sorts of incentives matter, and how, to individuals. Although the role of pecuniary incentives within firms has been long studied in the sociology and management literature, within economics, the stream of work has its roots in contract theory. The basic questions for economists then revolve around how workers respond to incentives and the optimal design of those incentives. Early empirical work used personnel data to measure the effects of compensation on individual productivity levels, but a difficult econometric challenge arose as most (all?) observed incentive contracts in naturally occurring settings are endogenous, making causality difficult to establish. Economic experiments introduce exogeneity in incentives that are by design orthogonal to other management practices, opening the possibility of identifying the causal relationship between pecuniary incentives and effort levels of individual workers. We take this literature as a starting point to discuss work on how market design can be used to improve the workplace, with each of the bins showcased by laboratory and field experiments.

5.1. Putting Market Design to Work to Increase Effort in the Workplace

Although designing individual incentive schemes, such as clawback bonus contracts, have proven quite fruitful, market design comes in many flavors. One such example is rewarding employees based on relative payoffs (tournaments), which are ubiquitous in the workplace. Job promotions, earning bonuses, employee of the month, and the like all revolve around relative assessment of workers. The literature on the market design of tournaments has witnessed deep contributions theoretically, showcasing the benefits and costs of relative performance incentives (Lazear and Rosen 1981, Holmstrom 1982). However, empirically testing the theoretical predictions has proven quite difficult, as observing individual effort, the main outcome variable in the theory, has been elusive in field settings.

One focus, therefore, has been to use laboratory experiments that are able to measure effort directly. An early experiment in this spirit is from Bull et al. (1987). They design a laboratory experiment with student subjects to explore the first-order predictions from tournament models concerning effort provision. In practice, the object of choice in the experiment is asking the student to circle a number that represents their effort choice, with higher numbers yielding a better chance of winning but being more costly in a convex manner. This effort choice is then added to an idiosyncratic shock, or luck component, that is uniformly distributed, and their sum determines the tournament outcome.

This approach, importantly and cleverly, embeds the essential elements of the tournament theory in the effort choice setting. Bull et al. (1987) find that, although there is considerable noise in individual play, effort levels converge to theoretical predictions in aggregate. That is, although individual effort level choices in the laboratory are quite noisily distributed around the equilibrium prediction, in aggregate, the theory performs remarkably well: a key win for the theory.

Following Bull et al. (1987), subsequent work has explored beyond equilibrium play to examine how other factors, such as selection (Camerer and Lovallo 1999, Eriksson et al. 2009, Cason et al. 2010, Müller and Schotter 2010, Faravelli et al. 2015), feedback (Blanes and Nossol 2011), sabotage (Harbring and Irlenbusch 2008, Harbring and Irlenbusch 2011), dynamic tournaments (Casas-Arce and Martínez-Jerez 2009, Sheremeta 2010, Brown and Minor 2014, Liu et al. 2014, Mostagir et al. 2019), and sex (Gneezy et al. 2003, Dargnies 2012), affect individual effort level choices in tournaments.

In addition to the aforementioned factors, there is also a large experimental literature focusing on the effect of tournament size and prize structure. For instance, Sheremeta (2011) finds that the average effort per participant is lower in four player contests than in two player contests. Using experiments with different group sizes, Gneezy and Smorodinsky (2006) and Morgan et al. (2012) find further evidence that average individual effort decreases with the number of players. Conversely, Harbring and Irlenbusch (2003) examine rank order tournaments and find the average effort increasing with the number of players. Also examining rank order tournaments, Orrison et al. (2004) vary the fraction of winner and loser prizes in tournaments of different sizes and find no trend relating effort with the size of the contest when the luck term is uniformly distributed.

Further expanding this literature to field settings, List et al. (2020) use complementary laboratory and field experiments to analyze how the number of competitors, or the size of tournament, affects effort levels under different distributions of luck.[10] In the previous literature, the key assumption that the luck term is drawn from a uniform distribution makes this particular question moot when workers are risk neutral, because the number of entrants in this case does not affect equilibrium effort levels. However, there are several instances where such an assumption need not hold. Indeed, in important ways, endogenizing the number of players allowed to enter the tournament becomes an interesting market design consideration, as was seemingly anticipated by the English poet Milton (1628), who once quipped that "luck is the residue of design."[11]

Consider the thought experiment of a worker innovating on the job with one prize awarded to the best innovator. Suppose that there are many potentially successful innovation paths, and workers arbitrarily choose a path. Each worker then expects to be a successful innovator, but she also expects that at least one other worker will be successful too. Hence, effort is crucial in determining the winning innovator, and investment is high (and even higher if the number of competitors expands). Alternatively, suppose that workers believe the chance of developing a very successful product are small. In that case, they expect that luck will play a crucial role in selecting the winner. If there are many innovators, each worker knows that at least one of them will be lucky but also knows that it is unlikely that it will be them. Because luck is much more important in selecting the winner than effort in this case, workers invest little effort, instead relying on luck to determine the winner. This result is exacerbated as the number of competitors increases, leading to an even lower level of investment as the number of competitors increases. Several real-world examples abound, from development of autonomous vehicles to finding medicinal drug breakthroughs.

The theory of List et al. (2020) highlights these intuitions and shows that if the distribution of the uncertainty component is skewed, the number of competitors allowed in the competition has a critical influence on equilibrium effort levels. As the number of competitors increases, a worker's equilibrium effort level (i) decreases if there is small mass on good luck, (ii) remains constant if the luck component is drawn from a uniform density, and (iii) increases if there is a large mass on good luck. The intuition is that the marginal benefit from committing effort depends on both the number of competitors and on the good draw mass, which depends critically on the density function's slope.

The empirical approach of List et al. (2020) to test the theory begins in the laboratory and closely follows the approach of Bull et al. (1987). This permits a study of labor markets that differ only in the shape of the density function, allowing a unique insight into whether changes in the component's shape itself can lead to predicted behavioral changes. Their second empirical approach is to use a field experiment that resembles the important features of the theory but maintains enough control to allow a formal test of the theory. In doing so, it is important to create an experimental design that exogenously varies their major treatment variable—number of competitors—in an environment that permits an understanding of the other important features of the situation.

This is not simple because one needs to find a naturally occurring environment where the random stochastic component takes a shape that is well understood by the participants. List et al. (2020) ended up choosing recreational commercial fisherman, where the private ponds were stocked with rainbow trout or salmon trout. Importantly, the fishing pond permits a natural test of the theory for the case of a decreasing density function. This is because of the fish schooling: the density function of luck is decreasing because schools never cover more than just a few rectangles of where the competitors are placed, and hence the amount of mass on having good luck is quite small.

Overall, the laboratory results are in line with the tournament theory. Most importantly, they report that when exploring tournaments with two, three, and four players, the impact of group size on effort is consonant with theory. The field experimental results complement these insights by providing evidence consonant with the theory within a special case of the theory—when the density function is negatively skewed.

In this case, the author's report evidence that adding competitors decreases individual effort levels especially when they control for fatigue.

Beyond testing theory, the received results enhance the manager's choice set by showing that the number of competitors in a relative pay incentive scheme has important effects on individual effort levels. Methodologically, the study showcases the power of testing the theory using complementary laboratory and field experiments within the same study (rather than across studies as shown in the loss aversion experiments). First, using an artificial setting that permits an examination of markets that differ only in the shape of the density function allows a test of such effects that would be difficult to identify in naturally occurring data. Pairing that with a second empirical approach that maintains randomization, and closely resembles the important features of the theory, as a field test provides much stronger inference of the underlying data patterns than either the laboratory or field approach could provide in isolation.

5.2. Understanding Nonpecuniary Incentives in the Workplace: From Gift Exchange to Corporate Social Responsibility

Perhaps the least understood aspect of the designer's quiver is how nonpecuniary incentives work in labor markets. However, recent literature is beginning to provide insight into this fascinating area. A few examples include Bradler et al. (2016) and Gallus (2017), who use field experiments to explore the power of employee recognition on employee performance. The results are impressive; for example, the Wikipedia natural field experiment of Gallus (2017) shows that her symbolic awards have a sizeable effect on newcomer retention, which persists over a span of four quarters. Likewise, Cohn et al. (2015) reveal how the perceptions of the fairness of pay affects effort provision.

Relatedly, field experimental research into unconditional gifts in the workplace is a burgeoning area of research—see Gneezy and List's (2006) surprise pay raises, pay cuts in Kube et al. (2013), and in-kind gifts in Kube et al. (2012). Recent work on corporate social responsibility (CSR) highlights that gifts for the social good can also have important labor market effects (Tonin and Vlassopoulos 2015, List and Momeni 2020). In this section, we focus attention to these two strands of research: gift exchange and CSR.

5.2.1. Gift Exchange. A common result that is found in large data sets is that employers are observed paying above the market equilibrium wage. When effort is monitored, workers exert more than the minimum effort level. This empirical observation has induced a set of economic models based on the assumption of there being a positive relationship between worker wages and effort levels (Akerlof 1982, Akerlof and Yellen 1990). The equilibrium of these models is for employers to offer higher than market clearing wages and for workers to reciprocate with high effort levels, making the situation a win-win for principal and agent.

Within experimental economics, the literature on social preferences has become one of the most influential areas of research (Camerer and Weigelt 1988, Fehr et al. 1993, Levitt and List 2007, Cooper and Kagel 2016). Findings from such games have been interpreted as providing strong evidence that many agents behave in a reciprocal manner even when the behavior is costly and yields neither present nor future material rewards. That workers respond positively to employers who offer a generous wage has been heavily documented in the laboratory, suggesting that reciprocity itself can function as a powerful incentive for employers to deviate from purely selfish behavior (Fehr et al. 1997, Fehr and Falk 1999). As a result, reciprocity has been found to explain employer generosity and worker cooperation in the presence of incomplete contracts (Fehr and Falk 1999, Bolton and Ockenfels 2000). Even beyond one-shot laboratory games, reciprocity and repeated game incentives have been found to reinforce each other (Gächter and Falk 2001). Moreover, the potential for employers to again reciprocate worker generosity by positively or negatively altering worker payoffs has been found to increase workers' effort levels (Fehr et al. 1997). Fehr and Falk (2008) provide a comprehensive summary of studies that show a sustained deviation from equilibrium wages and effort levels in experimental markets.

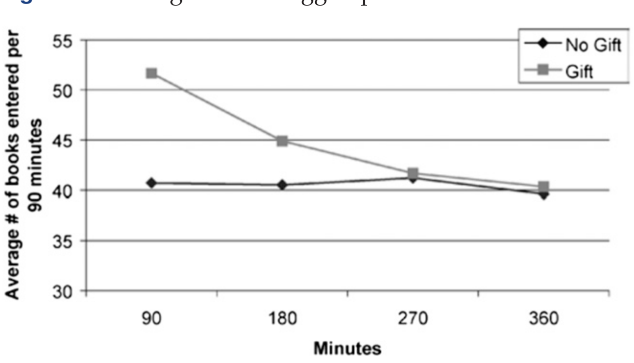
Ultimately, social preferences such as reciprocity can behave as a contract enforcement device that raises efficiency gains, and a well-designed market must carefully acknowledge the contextual influence of such forces (see Section 3). Fehr and Falk (2002) show that in addition to contract and principal-agent theory, nonpecuniary incentives significantly shape human behavior. In fact, policies that align financial incentives with social objectives have been found in some situations to lead to inefficient outcomes by crowding out other-regarding preferences (Bowles and Hwang 2008). Bowles and Gintis (2013) provide a taxonomy of cooperative and generous behavior not explained by conventional self-interest hypotheses. Similarly, various studies provide models and evidence of altruistic behavior largely explained by strong reciprocity, a predisposition to cooperate but also to punish those who fail to conform to the norms of cooperation (Bowles and Gintis 2004, Fehr and Fischbacher 2005, Gintis et al. 2008, Bowles and Gintis 2013, Mussweiler and Ockenfels 2013). Punishment and retaliation are further discussed as strong reciprocity enforcement devices (Fehr and Gächter 2000b, 2002; Fehr and Falk 2002).

These results have been widely applied outside the laboratory, becoming a descriptor of environments far removed from the domains of data generation. An early study exploring the importance of gift exchange in a naturally occurring market is List (2006), who used a series of field experiments in a product market to show that inference from these early games should be made with care because the environment might engender certain behaviors and that strategic reciprocity might masquerade as social preferences in certain instances.

Gneezy and List (2006) took this notion to the workplace by exploring the effects of gift exchange on worker productivity. They use two natural field experiments to explore gift exchange in the workplace. In the first, they recruited undergraduate students to participate in computerizing the holdings of a small library at the University of Chicago. In the no-gift treatment, individuals were offered a flat wage of $12 per hour. In the gift treatment, once the task was explained to participants, they were surprisingly paid $20 per hour rather than $12 per hour as advertised. The second field experiment was part of a door-to-door fundraising drive to support a university research center in North Carolina. Fundraising solicitors were recruited and told they would be paid $10 per hour, and those in the gift treatment were surprisingly told they would receive $20 per hour.

The main results from the library task are summarized in Figure 5. Two stark patterns are revealed in Figure 5. First, in line with earlier evidence from the laboratory, there are signs of significant gift exchange in the first few hours of the task. Second, there is a significant decrease in the gift exchange effect after a few hours, with no significant differences existing over a longer period. Importantly, the data reveal how the gift worked early on to induce higher output, but overall, the results show that with the same budget, the employer would have been better off paying market wages.[12]

Figure 5. Average Books Logged per Time Period



Source. Reproduced from Gneezy and List (2006).

Since this early work, the evidence on the efficacy of gift exchange in the workplace has come in various forms. Kube et al. (2012) find similar short run effect sizes as Gneezy and List (2006) using an in-kind gift. Likewise, using a negative wage gift, Kube et al. (2013) find a similarly large treatment effect when considering decreases in productivity. Alternatively, there are studies that report small and statistically weak gift exchange effects. For instance, Al-Ubaydli et al. (2015) report small effects despite that gifts used in their study were large: paying unskilled workers $18 an hour to pack envelopes. To reconcile the various facts around gift exchange in the workplace, Esteves-Sorenson (2018) carefully identifies several factors that could be underlying the inconsistent results. She concludes that "after dealing with all these confounds, our field test results are most consistent with a standard model: workers did not increase effort in response to fixed wage raises but did do so in response to a piece rate scheme" (p. 2). In this way, her results are quite consonant with the long-term estimates from Gneezy and List (2006). We return to the important question of inference from this body of work at the end of this section.

5.2.2. CSR. The role of CSR has been debated at least since Friedman famously described CSR as a *fundamentally subversive doctrine* in a 2007 *New York Times* article. In his usual combative style, Friedman described contrarians to his views on CSR as "puppets of the intellectual forces that have been undermining the basis of a free society" (Friedman 2007, p. 1). In this regard, the world has certainly turned its cheek to Friedman's advice, as today more than 90% of major businesses have specific programs dedicated to CSR. However, is this transformation a sign that the CSR business is a good business?

For their part, researchers have explored the efficacy of CSR in various venues. Broadly, most studies have focused on the demand side of the market, examining whether consumers are moved by the CSR programs of firms. Alternatively, only until recently has work begun to focus on the supply side to determine whether employees are affected by CSR. In this spirit, the notion of examining the supply side resembles gift exchange, except in this case, workers reciprocate higher effort when the firm does good for the society at large rather than for themselves directly.

Defined narrowly, thus far there exists little consensus on whether CSR investments positively impact the bottom line. Although some studies report a positive effect (Waddock and Graves 1997), others find mixed effects (Servaes and Tamayo 2013). As mentioned previously, however, one possible reason for these mixed findings is that, with few exceptions, empirical studies of CSR tend to focus primarily on the demand side of the market (Du et al. 2011).

In a study in *Management Science*, List and Momeni (2020) address this shortcoming in the literature by exploring the supply side effects of CSR within an online marketplace, with an emphasis placed on observing misbehaviors in the workplace. List and Momeni (2020) operationalized a test of CSR by conducting a natural field experiment using workers from Amazon's Mechanical Turk (MTurk). This approach has become a common one in the economics community searching for convenience labor market samples, because MTurk is an online labor market platform where ads are made to available workers. List and Momeni (2020) try to hire workers who land on their website after seeing an advertisement for work. On landing on their website, potential workers were randomized into one of the six treatments that had different wages and CSR language. In terms of the basic contracts, people were told that 10% of the total wage would be paid to workers upfront, and the remaining 90% would be paid after they completed the task. Accepting the contract without completing the task is one of two measures they use for worker misbehavior (because the worker is paid without delivery).

The task was for each worker to transcribe 10 images of short German texts, scanned from old German books. On average, each image was composed of around 30 words or 183 characters. The authors used German texts to make the task harder and less enjoyable. Workers who submitted all 10 images received the full wage specified in the contract. Before starting to work on any given image, workers were required to report if the image was legible. If an image was reported as unreadable, the worker skipped that image and moved on to the next, yet still received full pay. The second misbehavior naturally arises: Misreporting perfectly readable images as unreadable to avoid costly transcription was their second measure of misbehavior.

Although List and Momeni (2020) had six treatment conditions, the main treatment comparisons for our purposes were the outcomes of the baseline and a CSR treatment that had an advertisement that was identical to the baseline except included the following CSR message:

> Our firm is committed to give back in meaningful ways. We are passionate about encouraging education for the next generation. We do our part by donating money to influential non-profit organizations that support education for children from low socioeconomic backgrounds. In keeping with our philanthropic mission, we donate the equivalent of x% of our wage bill in cash (on behalf of all workers who help us with this project) to UNICEF Education Programs. UNICEF works tirelessly to ensure that every child regardless of gender, ethnicity or circumstances has access to a quality education. You may find out more about UNICEF Education Programs at: UNICEF.

The reported treatment effects are interesting, but one stark set of results stands out: The firm's use of CSR increased worker misbehavior. More specifically, workers who received a CSR message were more likely to become cheaters and cheated more often than those who were not incentivized with CSR. The data pattern is consistent with a *moral-licensing* impact: Doing good on one dimension (CSR work) allows the worker to shirk on another (misbehaving). Such an impact was anticipated by Bénabou and Tirole (2010), who note that "people who have recently done good in one dimension may feel immunized against negative (social or self) inferences, and thus later on act less morally constrained" (p. 6). The results raise the potential that, although CSR could very well have positive selection effects, there is a dark side of CSR that should be understood. More work is necessary.

5.3 Leveraging Behavioral Economics to Get More for Your Money in the Workplace

Behavioral economics has become much more than academic curiosity. Today, organizations as distinct as governments and firms use behavioral insights to chart a course of action. Although sister disciplines as varied as sociology, biology, and computer sciences have lent insight into the economic explorations, it is fair to say that, to date, psychology has made the deepest inroads in the behavioral economic revolution. This is because of the piercing nature of the received insight. For example, one of the deepest economic tenets—the basic independence assumption—has been under attack since the early experimental findings from the laboratory and field suggested that preferences are a function of current entitlements (Kahneman et al. 1990, List 2003). The most accepted theory explaining such behavior is broadly termed loss aversion.

One recent example leveraging loss aversion in the workplace is an eight-week-long field experiment from Hossain and List (2012). They explore whether worker productivity can be affected using simple loss averse framing of bonuses. In this manner, the treatment is particularly passive in comparison with previous field experiments that manipulated real endowments and explored choices (List 2004, 2011; Engelmann and Hollard 2010). For instance, in the main treatments, workers in the field experiment of Hossain and List (2012) received letters in the mail announcing the treatment. In the clawback treatment, for example, rather than giving the employees the bonus money before the work week commenced, it was given provisionally, where the relevant portion of the letter read as follows:

> "for every week in which the weekly production average of your team is below 400 units/hour, the salary enhancement will be reduced by RMB 80...."

Alternatively, in the standard bonus treatment, the description read as follows:

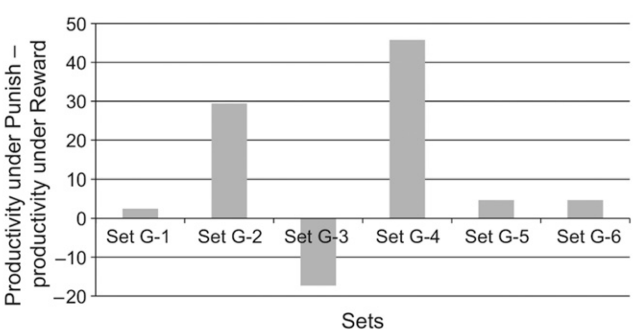
> "you will receive an RMB 80 bonus for every week the weekly production average of your team is above or equal to 400 units/hour. . ."

The setting in Hossain and List (2012) was a Chinese manufacturing plant that produced consumer electronics. The treatments were performed over both individual and team production (these passages are for the team setting). Hossain and List (2012) found that bonuses work: Posed as either gains or losses, workers in both teams and individually increased productivity when they received a bonus. More importantly for our purposes, they also find that teams and individuals respond more to bonuses posed as losses than as comparable bonuses posed as gains. Figure 6 summarizes empirical results from the six team sets in the field experiment of Hossain and List (2012). Importantly, Figure 6 shows that in five of six sets, the clawback treatment outperformed the standard bonus treatment. Overall, team productivity increased by 1% purely because of the framing manipulation. Comparable effects from the individual treatments, reported in the bottom panel of Figure 6, are of the same sign but are not statistically significantly different to each other. Although Figure 6 does not reveal the temporal treatment trends in the data, importantly when considering the treatment effects over time, the incentive effect does not wane: over the six-month study period, the loss-averse effects remain in the data. Furthermore, these productivity enhancements do not come with concomitant negative effects on the quality of work as measured by defect rates.

Subsequent work has largely replicated these results (Fryer et al. 2012, Levitt et al. 2016, Imas et al. 2017, Bulte et al. 2019). Like Hossain and List (2012), the main innovation in this literature is that agents receive an upfront payment that they must return in case their productivity or output fails to meet a certain threshold (i.e., a bonus scheme with upfront payments). Fryer et al. (2012), for example, provide pecuniary incentives to grammar and high school teachers to increase productivity as measured by the performance of their students. Although standard bonuses fail to increase teacher performance, leveraging the clawback scheme is quite effective.[13]

On a practical note, these examples highlight the interplay between pecuniary and nonpecuniary aspects of compensation in that this set of results shows that, conditional on bonuses being provided, framing matters. Given that framing can be adjusted costlessly, these approaches are simple ways in which firms can deepen the effects of pecuniary incentives. Theoretically, these results from the field provide an example consonant with loss aversion in a natural labor market setting. As such, the results provide concrete evidence of the generalizability of such insights from laboratory evidence and provide a natural example highlighting the complementarity of laboratory and field experiments.

Figure 6. Aggregate Differences in Per-Hour Productivities Under Punishment and Reward Treatments for Groups



Source. Reproduced from Hossain and List (2012).

5.4. Future Directions: Welfare, Structural Models, and Opening the Black Box of the Firm

In summarizing the range of work discussed in this section, several issues might have arisen to the astute reader. Early on, one might have asked, although these clawback incentives in Hossain and List (2012) are neat, would any worker like them in practice? That is, would the firm implementing such incentives soon find itself employee-less? Or, perhaps less extreme, maybe workers will remain in the firm but be miserable. Likewise, in considering the effects of gift exchange, one might wonder if the underlying motivations at work for the higher effort observed are because of altruism, warm glow, reciprocity, or some other consideration. Furthermore, in terms of market design choices, the reader might have wondered how CSR incentives affect other aspects of sorting, such as whether more productive workers prefer to be employed by firms that have a viable CSR program in place. Do workers view CSR like other workplace amenities, and what are the overall welfare effects of CSR?

These, and related questions, are important challenges that future work should take on, but the approach need not be blind, becaue some first steps have started to tackle these challenges. For example, one natural question that arises when the literature produces an incentive regime that improves performance is whether such a mechanism can be viewed as improving overall welfare levels. Considering the clawback scheme that leverages worker loss aversion, the evidence in Imas et al. (2017) and Bulte et al. (2019) suggests that workers value the clawback as a commitment device and therefore reciprocate by improving their performance in subsequent interactions. This result suggests that firms will not invoke anger or lose their employees. Of course, more work is necessary to be sure of such preliminary results, but this gives some sense of optimism that the loss aversion nudge will push workers to other firms. In this way, a call for future research is to identify the overall welfare effects of nudges.

In terms of future research paths on gift exchange literature, there is little field evidence about the nature of the observed social preferences toward employers. For instance, are workers acting on purely altruistic motives as in Becker (1974). Perhaps instead, workers are reacting to their warm glow in the spirit of Andreoni (1989). Or, perhaps the actual model at work is the original gift exchange view of Akerlof (1982). The importance of parsing these underpinnings is not just academic curiosity, because key market design considerations critically rely on whether extra worker effort is caused by enhanced value to the employer, pure value to self, or is strategically reciprocal in nature. However, there are many challenges to parsing such models.

In this case, a structural model combined with field experimental variation can provide insight. This is exactly the roadmap provided by DellaVigna et al. (2016), who combine a structural model with a closely linked field experiment to explore the lessons learned from the literature. What immediately falls out of such an exercise is that the extant literature lacks the key design features to parse the various interpretations. Two elements are missing from the designs in the literature: (i) there is no specification of the value to the employer of the worker's effort and (ii) a key unobservable is the cost of effort.

DellaVigna et al. (2016) design a field experiment to address both issues, yielding insight into the underlying primitives of workers. Their data and theoretical framework suggest that the warm glow model is key in explaining the received data in their setting. Importantly, such insight would have been difficult to obtain without the theoretical/experimental link offered in the study to permit a firmer understanding of the forces at work.

Hedblom et al. (2019) provide a similar contribution to the CSR literature, where they combine a structural approach with a natural field experiment to consider how CSR affects sorting into the workplace. Using data from more than 1,000 job seekers, they report strong evidence on the efficacy of CSR: it attracts employees who are more productive, produce higher quality work, and have more highly valued leisure time. In terms of enhancing the labor pool, use of CSR increases the number of applicants by 25%, an impact comparable to the effect of a 36% increase in wages. This work balances the moral hazard effects reported in List and Momeni (2020) and showcases the key complementarities in CSR and pecuniary incentives. In addition, this work provides compelling reasons for why firms might create and actively engage in CSR activities. We view these two examples as representing a key call for future work that combines the structural and field experimental approaches to shed new insights into both old and new areas of study.

This line of work underscores the import of leveraging firms and opening the black box of the firm. As Levitt and List (2009) discuss, there have been three distinct waves of field experimental research in economics. We are currently in the third wave, which has brought a deeper and broader exploration of economic phenomena in organizations. Within this generation of field experiments, an attractive approach is creating partnerships with private entities. Low hanging fruit remain in the areas of optimal worker incentive schemes, workplace design, wellness and health programs, and several related topics in the black box of the firm. One such area in the black box relates to a longstanding puzzle in economics: the striking differences in firm-level productivity across space and time. With total factor productivity ratios about 3:1 across high productivity and low productivity (90th percentile to 10th percentile) firms, understanding their sources is of first-order import. One such line of work considers management practice. Recently, rich literature has developed that provides key evidence around management's role in such disparities (Bloom and Van Reenen 2007, 2011; Gosnell et al. 2019). Much work remains, and we envisage an active area for future work for decades to come.

6. Concluding Remarks

In the last few decades, *Management Science* has published foundational work in market design and human behavior, as outlined in our Introduction. Today, economists, computer scientists, operations researchers, managers, and others are increasingly asked to design mechanisms for markets and organizations. Many of the applications are motivated by failures of incentives in markets and organizations and the urgent need to understand and fix the design flaws. One lesson learned from these efforts, as selectively surveyed in our article, is that institutional details matter. Even small changes in the rules can have a dramatic impact on the effectiveness and efficiency of a market. For instance, whether an otherwise identical auction ends with a hard or a soft close can significantly affect bidding, revenues, and allocative efficiency. Market design forces researchers to pay attention to details that might otherwise be overlooked. Our survey illustrates how the practical lessons from market design activities in various contexts may accumulate to become broad and sound scientific knowledge, which in turn promotes better and more reliable markets and organizations.

Some critics sometimes complain that economic theory is too disconnected from practical problems. However, in all cases that we survey, game theory proved helpful to develop intuition and to address real-life challenges. That said, market design may require decisions that are beyond current knowledge. Part of the reason is that theory must abstract from institutional and behavioral real-world complexities. For instance, people often follow their own—bounded—rationality, characterized by limitations of motivation, cognition, and adaptation. Boundedly rational agents only have limited cognitive abilities and bounded willpower, which constrains optimally responding to auction and matching mechanisms (Ockenfels and Selten 2005, Engelbrecht-Wiggans and Katok 2008, Hassidim et al. 2016). Bounded self-interest implies that humans are often willing to sacrifice their own interests to help others, which is key for understanding how to engineer trust in the gig economy (see Section 3). This is why practical market design is often fruitfully complemented by laboratory and field experiments that test game theory's predictions and provide a testbed and proof of concept before introducing new mechanisms into operating markets. A mechanism that works fine under simplifying assumptions about human behavior may fail under descriptively more relevant assumptions.

Our survey includes discussion of where we find gaps in our knowledge and where we believe more research is promising for auction, matching, feedback, and labor market design. There is also more general insight that holds across market design domains. In some cases, for instance, markets with theoretically attractive properties involve transactions that are perceived by many as repugnant. This can be an important constraint on market design. For instance, buying and selling kidneys for transplantation or trading school and university admission is illegal in most countries (Roth 2007). Thus, a stream of recent studies is concerned with understanding the empirical nature and robustness of such constraints to reconcile ethical concerns with economic effectiveness (Leider and Roth 2010; Ambuehl et al. 2015, 2019a, b; Kirchler et al. 2016; Ambuehl 2017).

Many opportunities and challenges in market design have to do with recent advances in computer and communication technology, which often allow for radical innovation in market design. Indeed, smart markets are popping up everywhere, from new kidney exchanges, dating, job and ride hailing markets, ad and spectrum auctions, to innovative climate, electricity, and financial markets. The development of these markets not only creates new business opportunities to benefit our social and economic lives but also improve our scientific understanding of engineering incentives and markets. There is probably no other field in economics and management science where researchers and practitioners gain so much by carefully listening to and working with one another. In this spirit, perhaps the most foundational change for generation of knowledge is that researchers will increasingly have to use the carpool lane in their own work, because riding alone will soon be an inefficient choice in the knowledge production game.

Acknowledgments

The authors thank the editor-in-chief, the associate editor, and three anonymous reviewers for their constructive comments, which significantly improved the paper.

Endnotes

[1] This section is an adjusted and updated, and much shortened, version of the account of Ockenfels and Roth (2013) of the literature on sniping in auctions for consumer goods. To simplify our exposition, we will use the present tense when we talk about Amazon auctions in this section, although Amazon shut down its auction platform many years ago.

[2] This section is mostly based on Cramton (2013). Because of space restrictions, we cannot discuss issues related to package bidding here, although they are an important part of the auction design literature (Cramton et al. 2006, Milgrom 2017), with important papers published in *Management Science*, such as Pekec and Rothkopf (2003), Kwasnica et al. (2005), and Goetzendorff et al. (2015). In this context, some early papers also addressed strategic timing, such as Rassenti et al. (1982) and Banks et al. (1989).

[3] Klemperer (2020) proposes an ending rule for spectrum auctions that is somewhat closer to eBay's hard close, namely a hybrid of the ascending, soft-close auction format and the sealed-bid format, which he calls *Anglo-Dutch*. The idea is that the early bidding is like in the simultaneous ascending auction, but bidders can make final sealed bids at the end of the auction. Klemperer argues that this kind of hard close can discourage collusion in the dynamic phase of the auction, because the last-minute round allows bidders to renege on any deals without fear of retaliation, and, because the final bids induce some uncertainty about the winner, this can also attract entrants. Such concerns are not relevant for the choice of eBay's hard close ending rule in their single-object auctions, however.

[4] Studies on eBay reveal that bidders do not bid truthfully early in the auction but that much of the price discovery is done only in the closing seconds of the auction (Ariely et al. 2005).

[5] Here, each lot corresponds to a specific quantity of spectrum, measured in either MHzPop or in eligibility points. The bidder starts with an initial eligibility based on the bidder's initial deposit. To maintain this level of eligibility in future rounds, the bidder needs to bid on a sufficiently large quantity of spectrum in the current round, where *sufficiently large* is stated as some percentage, typically between 80% and 100% of the bidder's current eligibility. If the bidder bids on a smaller quantity, the bidder's eligibility is reduced in future rounds.

[6] Ely and Hossain (2009) suggest from their field experiment that the availability of closely substitutable auctions on eBay may reduce the overall benefit of sniping.

[7] Another auction context where much has been learned from laboratory human subject research is the practical design of procurement auctions, with much research published in *Management Science*

(Katok and Roth 2004; Engelbrecht-Wiggans and Katok 2008; Davis et al. 2011, 2014; Chaturvedi et al. 2014; Fugger et al. 2015).

[8] This section is based on, and partly taken from, the school choice literature review in Chen et al. (2018) and Chen and Kesten (2019).

[9] This section is mostly based on, and partly taken from, Chen and Kesten (2017, 2019).

[10] In a set of innovative papers, Boudreau et al. (2011) examine naturally occurring data to explore the effects of group sizes in tournaments on software development. Their variable of interest is the score assigned to a solution of a software problem rather than effort levels.

[11] Many scholars credit English poet John Milton(1608-1674) for this quote—specifically “At a Vacation Exercise in the College” (1628)—but the saying does not seem to appear in any of Milton’s writings. Branch Rickey, a Major League Baseball Executive is also credited with making this statement in 1915.

[12] Building on Gneezy and List (2006), Ockenfels et al. (2015b) and Sliwka and Werner (2017) have explored the timing of wage increases in more detail showing that performance can be raised when providing smaller but more frequent wage increases. As Sliwka and Werner argue, these patterns (as well as the pattern detected by Gneezy and List 2006) can be well organized in a simple dynamic model of reciprocity, where workers reciprocate higher wages but adapt their reference points over time. In turn, reciprocal reactions to wage increases naturally wear off but can be made more persistent when the wages are increased gradually.

[13] Further evidence for the importance of loss aversion in the design of compensations schemes in field settings is provided by Ockenfels et al. (2015a), who study the bonus scheme of a multinational company. The bonus scheme created a clear reference point as managers in parts of the company learned not only the size of their bonus but also the percentage share of their bonus relative to their direct colleagues. Studying the association between these bonus shares and job satisfaction, Ockenfels et al. (2015a) find a substantial asymmetry around this 100% threshold, well in line with the patterns predicted by loss aversion.

References

Abdulkadiroğlu A, Sönmez T (2003) School choice: A mechanism Design approach. *Amer. Econom. Rev.* 93(3):729–747.
Abdulkadiroğlu A, Che YK, Yasuda Y (2011) Resolving conflicting preferences in school choice: The “Boston Mechanism” reconsidered. *Amer. Econom. Rev.* 101(1):399–410.
Abdulkadiroğlu A, Pathak PA, Roth AE (2005a) The New York City high school match. *Amer. Econom. Rev.* 95(2):364–367.
Abdulkadiroğlu A, Pathak PA, Roth AE, Sönmez T (2005b) The Boston public school match. *Amer. Econom. Rev.* 95(2):368–371.
Abdulkadiroğlu A, Che YK, Pathak PA, Roth AE, Tercieux O (2017) *Minimizing Justified Envy in School Choice: The Design of New Orleans’ OneApp* (National Bureau of Economic Research, Cambridge, MA).
Akerlof GA (1982) Labor contracts as partial gift Exchange. *Quart. J. Econom.* 97(4):543–569.
Akerlof GA, Yellen JL (1990) The fair wage-effort hypothesis and unemployment. *Quart. J. Econom.* 105(2):255–283.
Aldrich EM, Friedman D (2018) Order protection through delayed messaging. Working paper, University of California, Santa Cruz.
Aldrich EM, López Vargas K (2019) Experiments in high-frequency trading: Comparing two market institutions. *Exp. Econom.* https://doi.org/10.1007/s10683-019-09605-2.
Aldrich EM, Grundfest J, Laughlin G (2016) The flash crash: A new deconstruction. Working paper, University of California, Santa Cruz.
Al-Ubaydli O, Andersen S, Gneezy U, List JA (2015) Carrots that look like sticks: Toward an understanding of multitasking incentive schemes. *Southern Econom. J.* 81(3):538–561.
Ambrus A, Greiner B (2012) Imperfect public monitoring with costly punishment: An experimental study. *Amer. Econom. Rev.* 102(7):3317–3332.
Ambuehl S (2017) An offer you can’t refuse? Incentives change how we inform ourselves and what we believe, Working paper, Cesifo No. 6296
Ambuehl S, Bernheim BD, Ockenfels A (2019a) Projective paternalism. Working paper, University of Cologne, Germany.
Ambuehl S, Niederle M, Roth AE (2015) More money, more problems? Can high pay be coercive and repugnant? *Amer. Econom. Rev.* 105(5):357–360.
Ambuehl S, Ockenfels A, Stewart C (2019b) Attention and selection effects. Working paper, Rotman School of Management.
Andreoni J (1989) Giving with impure altruism: Applications to charity and Ricardian equivalence. *J. Political Econom.* 97(6):1447–1458.
Anwar S, McMillan R, Zheng M (2006) Bidding behavior in competing auctions: Evidence from eBay. *Eur. Econom. Rev.* 50(2):307–322.
Aquilina M, Budish E, O’Neill P (2020) Quantifying the high-frequency trading “arms race”: A simple new methodology and estimates. United Kingdom Financial Conduct Authority Occasional Paper 50, Financial Conduct Authority, London.
Ariely D, Ockenfels A, Roth A (2005) An experimental analysis of ending rules in Internet auctions. *RAND J. Econom.* 35(4):890–907.
Arrow KJ, Forsythe R, Gorham M, Hahn R, Hanson R, Ledyard JO, Levmore S, et al (2008) The promise of prediction markets. *Science* 320(5878):877–878.
Athey S, Ellison G (2011) Position auctions with consumer search. *Quart. J. Econom.* 126(3):1213–1270.
Ausubel LM, Baranov O (2020) Revealed preference and activity rules in dynamic auctions. *Internat. Econom. Rev.* 61(2):471–502.
Ausubel LM, Cramton P, Milgrom P (2006) The clock-proxy auction: A practical combinatorial auction design. Cramton P, Shoham Y, Steinberg R, eds. *Combinatorial Auctions* (MIT Press, Cambridge, MA), 115–138.
Backus M, Blake T, Masterov DV, Tadelis S (2015) Is sniping a problem for online auction markets? *Proc. 24th Internat. Conf. World Wide Web* (ACM, New York), 88–96.
Bajari P, Hortaçsu A (2004) Economic insights from Internet auctions. *J. Econom. Literature* 42(2):457–486.
Balafoutas L, Beck A, Kerschbamer R, Sutter M (2013) What drives taxi drivers? A field experiment on fraud in a market for credence goods. *Rev. Econom. Stud.* 80(3):876–891.
Balafoutas L, Kerschbamer R, Sutter M (2017) Second-degree moral hazard in a real-world credence goods market. *Econom. J. (London)* 127(599):1–18.
Balafoutas L, Nikiforakis N, Rockenbach B (2014) Direct and indirect punishment among strangers in the field. *Proc. Natl. Acad. Sci. USA* 111(45):15924–15927.
Banks J, Ledyard J, Porter D (1989) Allocating uncertain and unresponsive resources: An experimental approach. *RAND J. Econom.* 20(1):1–25.
Bar-Isaac H, Tadelis S (2008) Seller reputation. *Foundations Trends Microeconom.* 4(4):273–351.
Becker GS (1974) A theory of social interactions. *J. Political Econom.* 82(6):1063–1093.
Bénabou R, Tirole J (2010) Individual and corporate social responsibility. *Economica* 77(305):1–19.
Benartzi S, Thaler RH (1995) Myopic loss aversion and the equity premium puzzle. *Quart. J. Econom.* 110(1):73–92.
Benos E, Sagade S (2016) Price discovery and the cross-section of high-frequency trading. *J. Financial Markets* 54–77.
Benos E, Brugler J, Hjalmarsson E, Zikes F (2017) Interactions among high-frequency traders. *J. Financial Quantitative Analysis* 52(4): 1375–1402.
Bichler M, Goeree JK, eds. (2017) *Handbook of Spectrum Auction Design*, 1st ed. (Cambridge University Press, Cambridge, UK).
Bichler M, Gupta A, Ketter W (2010) Research commentary—Designing smart markets. *Informs Systems Res.* 21(4):688–699.

Bierbrauer F, Ockenfels A, Pollak A, Rückert D (2017) Robust mechanism design and social preferences. *J. Public Econom.* 149(C):59–80.

Blanes VJ, Nossol M (2011) Tournaments without prizes: Evidence from personnel records. *Management Sci.* 57(10):1721–1736.

Bloom N, Van Reenen J (2007) Measuring and explaining management practices across firms and countries. *Quart. J. Econom.* 122(4):1351–1408.

Bloom N, Van Reenen J (2011) Human resource management and productivity. Card D, Ashenfelter O, eds. *Handbook of Labor Economics* (Elsevier, New York), 1697–1767.

Bó IL, Hakimov R (2020) Iterative vs. standard deferred acceptance: Experimental evidence. *Econom. J. (London)* 130(626):356–392.

Bolton GE, Ockenfels A (2000) ERC—A theory of equity, reciprocity and competition. *Amer. Econom. Rev.* 90(1):166–193.

Bolton GE, Ockenfels A (2009) The limits of trust in economic transactions: Investigations of perfect reputation systems. Cook KS, Snijders S, Buskens V, Cheshire C, eds. *eTrust: Forming Relationships in the Online World* (Russel Sage, New York), 15–36.

Bolton GE, Ockenfels A (2012) Behavioral economic engineering. *J. Econom. Psychol.* 33(3):665–676.

Bolton GE, Ockenfels A (2014) Does laboratory trading mirror behavior in real world markets? Fair bargaining and competitive bidding on eBay. *J. Econom. Behavioral Organizations* 97(C):143–154.

Bolton GE, Greiner B, Ockenfels A (2013) Engineering trust: Reciprocity in the production of reputation information. *Management Sci.* 59(2):265–285.

Bolton GE, Greiner B, Ockenfels A (2018) Dispute resolution or escalation? The strategic gaming of feedback withdrawal options in online markets. *Management Sci.* 64(9):4009–4031.

Bolton GE, Katok E, Ockenfels A (2004) How effective are electronic reputation mechanisms? An experimental investigation. *Management Sci.* 50(11):1587–1602.

Bolton GE, Katok E, Ockenfels A (2005) Cooperation among strangers with limited information about reputation. *J. Public Econom.* 89(8):1457–1468.

Bolton GE, Kusterer DJ, Mans J (2019) Inflated reputations: Uncertainty, leniency, and moral wiggle room in trader feedback systems. *Management Sci.* 65(11):5371–5391.

Bolton GE, Mans J, Ockenfels A (2020a) Norm enforcement in markets: Group identity and the volunteering of feedback. *Econom. J. (London)* 130(629):1248–1261.

Bolton GE, Ockenfels A, Ebeling F (2011) Information value and externalities in reputation building. *Internat. J. Industrial Organizations* 29(1):23–33.

Bolton GE, Breuer K, Greiner B, Ockenfels A (2020b) Fixing feedback revision rules in online markets. Working paper, Vienna University of Economics and Business, Vienna.

Boudreau KJ, Lacetera N, Lakhani KR (2011) Incentives and problem uncertainty in innovation contests: An empirical analysis. *Management Sci.* 57(5):843–863.

Bowles S, Gintis H (2004) The evolution of strong reciprocity: cooperation in heterogeneous populations. *Theoretical Population Biol.* 65(1):17–28.

Bowles S, Gintis H (2013) *A Cooperative Species: Human Reciprocity and Its Evolution* (Princeton University Press).

Bowles S, Hwang SH (2008) Social preferences and public economics: Mechanism design when social preferences depend on incentives. *J. Public Econom.* 92(8-9):1811–1820.

Bradler C, Dur R, Neckermann S, Non A (2016) Employee recognition and performance: A field experiment. *Management Sci.* 62(11):3085–3099.

Brogaard J, Hendershott T, Riordan R (2014) High-frequency trading and price discovery. *Rev. Financial Stud.* 27(8):2267–2306.

Brown J, Minor DB (2014) Selecting the best? Spillover and shadows in elimination tournaments. *Management Sci.* 60(12):3087–3102.

Brown J, Morgan J (2009) How much is a dollar worth? Tipping vs. equilibrium coexistence on competing online auction sites. *J. Political Econom.* 117(4):668–700.

Budish E (2011) The combinatorial assignment problem: Approximate competitive equilibrium from equal incomes. *J. Political Econom.* 119(6):1061–1103.

Budish E, Cantillon E (2012) The multi-unit assignment problem: Theory and evidence from course allocation at Harvard. *Amer. Econom. Rev.* 102(5):2237–2271.

Budish E, Kessler JB (2019) Bringing real market participants' real preferences into the laboratory: An experiment that changed the course allocation mechanism at Wharton. Working paper, NBER, Cambridge, MA.

Budish E, Cramton P, Shim J (2015) The high-frequency trading arms race: Frequent batch auctions as a market design response. *Quart. J. Econom.* 130(4):1547–1621.

Budish E, Lee RS, Shim JJ (2019) Will the market fix the market? A theory of stock exchange competition and innovation. Working paper, NBER, Cambridge, MA.

Budish E, Cachon GP, Kessler JB, Othman A (2017) Course match: A large-scale implementation of approximate competitive equilibrium from equal incomes for combinatorial allocation. *Oper. Res.* 65(2):314–336.

Budish E, Cramton P, Kyle AS, Lee J, Malec D (2020) Flow trading. Working paper, University of Cologne, Germany.

Bull C, Schotter A, Weigelt K (1987) Tournaments and piece rates: An experimental study. *J. Political Econom.* 95(1):1–33.

Bulte E, List JA, van Soest D (2019) Toward an understanding of the welfare effects of nudges: Evidence from a field experiment in Uganda. Working paper, NBER, Cambridge, MA.

Burtch G, Hong Y, Bapna R, Griskevicius V (2018) Stimulating online reviews by combining financial incentives and social norms. *Management Sci.* 64(5):2065–2082.

Cabral L, Li L (2015) A dollar for your thoughts: Feedback-conditional rebates on eBay. *Management Sci.* 61(9):2052–2063.

Camerer C, Lovallo D (1999) Overconfidence and excess entry: An experimental approach. *Amer. Econom. Rev.* 89(1):306–318.

Camerer C, Weigelt K (1988) Experimental tests of a sequential equilibrium reputation model. *Econometrica* 56(1):1–36.

Cao W, Sha Q, Yao Z, Gu D, Shao X (2019) Sniping in soft-close online auctions: Empirical evidence from Overstock. *Marketing Lett.* 30(2):179–191.

Carpenter J, Holmes J, Matthews PH (2011) Jumping and sniping at the silents: Does it matter for charities? *J. Public Econom.* 95(5): 395–402.

Casas-Arce P, Martínez-Jerez FA (2009) Relative performance compensation, contests, and dynamic incentives. *Management Sci.* 55(8):1306–1320.

Cason TN, Masters WA, Sheremeta RM (2010) Entry into winner-take-all and proportional-prize contests: An experimental study. *J. Public Econom.* 94(9-10):604–611.

Cassady R (1967) *Auctions and Auctioneering* (California University Press, Berkeley.).

Catalini C, Gans JS (2019) Some simple economics of the blockchain. Working paper, NBER, Cambridge, MA.

Chaturvedi A, Beil DR, Martínez-de-Albéniz V (2014) Split-award auctions for supplier retention. *Management Sci.* 60(7):1719–1737.

Chen Y, He Y (2019) Information acquisition and provision in school choice: A theoretical analysis. Working paper, University of Michigan, Ann Arbor.

Chen Y, He Y (2020) Information acquisition and provision in school choice: An experimental study. Working paper, University of Michigan, Ann Arbor.

Chen Y, Kesten O (2017) Chinese college admissions and school choice reforms: A theoretical analysis. *J. Political Econom.* 125(1): 99–139.

Chen Y, Kesten O (2019) Chinese college admissions and school choice reforms: An experimental study. *Games Econom. Behaviors* 115:83–100.

Chen Y, Li SX (2009) Group identity and social preferences. *Amer. Econom. Rev.* 99(1):431–457.

Chen Y, Sönmez T (2006) School choice: An experimental study. *J. Econom. Theory*. 127(1):202–231.

Chen Y, Harper FM, Konstan J, Li SX (2010) Social comparisons and contributions to online communities: A field experiment on MovieLens. *Amer. Econom. Rev.* 100(4):1358–1398.

Chen Y, Farzan R, Kraut R, YeckehZaare I, Zhang AF (2019) Motivating expert contributions to public goods: A personalized field experiment on Wikipedia. Working paper, University of Michigan, Ann Arbor.

Chen Y, Jiang M, Kesten O, Robin S, Zhu M (2018) Matching in the large: An experimental study. *Games Econom. Behaviors* 110:295–317.

Clarke EH (1971) Multipart pricing of public goods. *Public Choice* 11:17–33.

Cohn A, Fehr E, Goette L (2015) Fair wages and effort provision: Combining evidence from a choice experiment and a field experiment. *Management Sci.* 61(8):1777–1794.

Cooper DJ, Kagel JH (2016) Other-regarding preferences: A selective survey of experimental results. Kagel JH, Roth AE, eds. *Handbook of Experimental Economics,* Vol. 2 (Princeton University Press), 217–289.

Cotton C (2009) Multiple bidding in auctions as bidders become confident of their private valuations. *Econom. Lett.* 104(3):148–150.

Cramton P (2013) Spectrum auction design. *Rev. Industrial Organizations* 42(2):161–190.

Cramton P (2017) Electricity market design. *Oxford Rev. Econom. Policy* 33(4):589–612.

Cramton P, Ockenfels A (2012) Economics and design of capacity markets for the power sector. *Zeitschrift Energiewirtschaft* 36(2):113–134.

Cramton P, Ockenfels A (2017) The German 4G spectrum auction: Design and behaviour. *Econom. J. (London)* 127(605):F305–F324.

Cramton P, Ockenfels A, Stoft S (2013) Capacity market fundamentals. *Econom. Energy Environmental Policy* 2(2):27–46.

Cramton P, Ockenfels A, Tirole J (2017a) Translating the collective climate goal into a common climate commitment. *Rev. Environmental Econom. Policy* 11(1):165–171.

Cramton P, Shoham Y, Steinberg R, eds. (2006) *Combinatorial Auctions* (MIT Press, Cambridge, MA).

Cramton P, MacKay DJC, Ockenfels A, Stoft S, eds. (2017b) *Global Carbon Pricing: The Path to Climate Cooperation* (MIT Press, Cambridge, MA).

Cui R, Li J, Zhang DJ (2020) Reducing discrimination with reviews in the sharing economy: Evidence from field experiments on Airbnb. *Management Sci.* 66(3):1071–1094.

Dargnies MP (2012) Men too sometimes shy away from competition: The case of team competition. *Management Sci.* 58(11): 1982–2000.

Davis AM, Katok E, Kwasnica AM (2011) Do auctioneers pick optimal reserve prices? *Management Sci.* 57(1):177–192.

Davis AM, Katok E, Kwasnica AM (2014) Should sellers prefer auctions? A laboratory comparison of auctions and sequential mechanisms. *Management Sci.* 60(4):990–1008.

de Haan M, Gautier P, Oosterbeek H, van der Klaauw B (2016) The performance of school assignment mechanisms in practice. Working paper, University of Oslo, Norway.

Dellarocas C (2003) The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management Sci.* 49(10):1407–1424.

Dellarocas C, Wood CA (2008) The sound of silence in online feedback: Estimating trading risks in the presence of reporting bias. *Management Sci.* 54(3):460–476.

DellaVigna S, List JA, Malmendier U, Rao G (2016) Estimating social preferences and gift exchange at work. Working paper, NBER, Cambridge, MA.

Dhami SS (2016) *The Foundations of Behavioral Economic Analysis*, 1st ed. (Oxford University Press, UK).

Du S, Bhattacharya CB, Sen S (2011) Corporate social responsibility and competitive advantage: Overcoming the trust barrier. *Management Sci.* 57(9):1528–1545.

Dubins LE, Freedman DA (1981) Machiavelli and the Gale-Shapley algorithm. *Amer. Math. Monthly* 88(7):485–494.

Dulleck U, Kerschbamer R, Sutter M (2011) The economics of credence goods: An experiment on the role of liability, verifiability, reputation, and competition. *Amer. Econom. Rev.* 101(2):526–555.

Dur U, Hammond RG, Morrill T (2018) Identifying the harm of manipulable school-choice mechanisms. *Amer. Econom. J. Econom. Policy* 10(1):187–213.

Dustan A, De Janvry A, Sadoulet E (2015) *Flourish or Fail? The Risky Reward of Elite High School Admission in Mexico City* (Dept. of Economics, Vanderbilt Univ., Nashville, TN).

Dynarski S, Michelmore K, Libassi CJ, Owen S 2018. Closing the gap: The effect of a targeted, tuition-free promise on college choices of high-achieving, low-income students. Working paper, NBER, Cambridge, MA.

Edelman B, Ostrovsky M (2007) Strategic bidder behavior in sponsored search auctions. *Decision Support Systems* 43(1):192–198.

Edelman B, Ostrovsky M, Schwarz M (2007) Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *Amer. Econom. Rev.* 97(1):242–259.

Elfenbein DW, McManus B (2010) Last-minute bidding in eBay charity auctions. *Econom. Lett.* 107(1):42–45.

Ely JC, Hossain T (2009) Sniping and squatting in auction markets. *Am. Econom. J. Microeconom.* 1(2):68–94.

Engel C (2014) Social preferences can make imperfect sanctions work: Evidence from a public good experiment. *J. Econom. Behaviors Organization* 108:343–353.

Engelberg J, Williams J (2009) eBay's proxy bidding: A license to shill. *J. Econom. Behaviors Organization* 72(1):509–526.

Engelbrecht-Wiggans R, Katok E (2008) Regret and feedback information in first-price sealed-bid auctions. *Management Sci.* 54(4):808–819.

Engelmann D, Hollard G (2010) Reconsidering the effect of market experience on the 'Endowment Effect.' *Econometrica* 78(6):2005–2019.

Ergin H, Sönmez T (2006) Games of school choice under the Boston mechanism. *J. Public Econom.* 90(1):215–237.

Eriksson T, Teyssier S, Villeval MC (2009) Self-selection and the efficiency of tournaments. *Econom. Inquiry* 47(3):530–548.

Esteves-Sorenson C (2018) Gift exchange in the workplace: Addressing the conflicting evidence with a careful test. *Management Sci.* 64(9):4365–4388.

Fack G, Grenet J, He Y (2019) Beyond truth-telling: Preference estimation with centralized school choice and college admissions. *Amer. Econom. Rev.* 109(4):1486–1529.

Faravelli M, Friesen L, Gangadharan L (2015) Selection, tournaments, and dishonesty. *J. Econom. Behaviors Organization* 110:160–175.

Fehr E, Falk A (1999) Wage rigidity in a competitive incomplete contract market. *J. Political Econom.* 107(1):106–134.

Fehr E, Falk A (2002) Psychological foundations of incentives. *Eur. Econom. Rev.* 46(4-5):687–724.

Fehr E, Falk A (2008). Reciprocity in experimental markets. Plott C, Smith V, eds. *Handbook of Experimental Economics Results*, vol. 1 (North Holland), 325–334.

Fehr E, Fischbacher U (2005) The economics of strong reciprocity Gintis H, Bowles S, Boyd R, Fehr E, eds. *Moral Sentiments and Material Interests: The Foundations for Cooperation in Economic life* (MIT Press, Cambridge, MA) 151–193.

Fehr E, Gächter S (2000a) Cooperation and punishment in public goods experiments. *Amer. Econom. Rev.* 90(4):980–994.

Fehr E, Gächter S (2000b) Fairness and retaliation: The economics of reciprocity. *J. Econom. Perspective* 14(3):159–181.

Fehr E, Gächter S (2002) Altruistic punishment in humans. *Nature* 415(6868):137–140.

Fehr E, Gächter S, Kirchsteiger G (1997) Reciprocity as a contract enforcement device: Experimental evidence. *Econometrica* 65(4):833–860.

Fehr E, Kirchsteiger G, Riedl A (1993) Does fairness prevent market clearing? An experimental investigation. *Quart. J. Econom.* 108(2):437–459.

Fradkin A, Grewal E, Holtz D (2020) Reciprocity in two-sided reputation systems: Evidence from an experiment on Airbnb. Working paper, Boston University.

Friedman D (2010) Laboratory financial markets. Durlauf SN, Blume LE, eds. *Behavioural and Experimental Economics* (Palgrave Macmillan UK, London), 7387–7393.

Friedman D, Rust J, eds. (1993) *The Double Auction Market: Institutions, Theories, and Evidence* (Addison-Wesley Pub. Co., Reading, MA).

Friedman M (2007) The social responsibility of business is to increase its profits. Zimmerli W, Richter K, Holzinger M, eds. *Corporate Ethics and Corporate Governance* (Springer, London).

Fryer J, Roland G, Levitt SD, List J, Sadoff S (2012) Enhancing the efficacy of teacher incentives through loss aversion: A field experiment. Working paper, NBER, Cambridge, MA.

Fugger N, Katok E, Wambach A (2015) Collusion in dynamic buyer-determined reverse auctions. *Management Sci.* 62(2):518–533.

Gächter S, Falk A (2001) Reputation or reciprocity? An experimental investigation. Working paper, CESifo Network.

Gale D, Shapley LS (1962) College admissions and the stability of marriage. *Amer. Math. Monthly* 120(5):386–391.

Gallus J (2017) Fostering public good contributions with symbolic awards: A large-scale natural field experiment at Wikipedia. *Management Sci.* 63(12):3999–4015.

Gintis H, Bowles S, Boyd R, Fehr E (2003) Explaining altruistic behavior in humans. *Evolution Human Behavior* 24(3):153–172.

Gintis H, Henrich J, Bowles S, Boyd R, Fehr E (2008) Strong reciprocity and the roots of human morality. *Social Justice Res.* 21(2):241–253.

Glover B, Raviv Y (2012) Revenue non-equivalence between auctions with soft and hard closing mechanisms: New evidence from Yahoo! *J. Econom. Behaviors Organization* 81(1):129–136.

Gneezy U, List JA (2006) Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments. *Econometrica* 74(5):1365–1384.

Gneezy U, Smorodinsky R (2006) All-pay auctions—An experimental study. *J. Econom. Behaviors Organization* 61(2):255–275.

Gneezy U, Niederle M, Rustichini A (2003) Performance in competitive environments: Gender differences. *Quart. J. Econom.* 118(3):1049–1074.

Goetzendorff A, Bichler M, Day B, Shabalin P (2015) Compact bid languages and core-pricing in large multi-object auctions. *Management Sci.* 61(7):1684–1703.

Gong B, Liang Y (2019) A dynamic college admission mechanism in inner Mongolia: Theory and experiment. Working paper, East China Normal University, Shanghai, China.

Gosnell GK, List JA, Metcalfe RD (2019) The impact of management practices on employee productivity: A field experiment with airline captains. *J. Political Econom.* 128(4):1195–1233.

Gray S, Reiley DH (2013) Measuring the benefits to sniping on eBay: Evidence from a field experiment. *J. Econom. Management* 9(2): 137–152.

Greiner B, Ockenfels A (2009) Vom Labor ins Feld: Die ökonomik des Vertrauens. Becker J, Deutschmann C, eds. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* (Sonderhefte, Wiesbaden, Germany), 219–242.

Greiner B, Ockenfels A, Sadrieh A (2012) Internet auctions. Peitz M, Waldfogel J, eds. *Oxford Handbook of the Digital Economy* (Oxford University Press, NY), 306–342.

Groves T (1973) Incentives in teams. *Econometrica* 41(4):617–631.

Gul F (1997) A Nobel prize for game theorists: The contributions of Harsanyi, Nash and Selten. *J. Econom. Perspective* 11(3):159–174.

Gutt D, Neumann J, Zimmermann S, Kundisch D, Chen J (2019) Design of review systems—A strategic instrument to shape online reviewing behavior and economic outcomes. *J. Strategic Informs. Systems* 28(2):104–117.

Hakimov R, Kübler D (2019) Experiments on matching markets: A survey. *WZB Discussion Paper*.

Harbring C, Irlenbusch B (2003) An experimental study on tournament design. *Labour Econom.* 10(4):443–464.

Harbring C, Irlenbusch B (2008) How many winners are good to have?: On tournaments with sabotage. *J. Econom. Behaviors Organization* 65(3-4):682–702.

Harbring C, Irlenbusch B (2011) Sabotage in tournaments: Evidence from a laboratory experiment. *Management Sci.* 57(4):611–627.

Harsanyi JC (1967) Games with incomplete information played by "Bayesian" players: Part I. *Management Sci.* 14(3):159–182.

Harsanyi JC (1968a) Games with incomplete information played by "Bayesian" players: Part II. *Management Sci.* 14(3):320–334.

Harsanyi JC (1968b) Games with incomplete information played by "Bayesian" players: Part III. *Management Sci.* 14(3):486–502.

Hassidim A, Romm A, Shorrer RI (2016) "Strategic" behavior in a strategy-proof environment. *Proc. 2016 ACM Conf. Econom. Comput, Maastricht, Netherlands, July, 2016* (ACM, New York), 763–764.

Hastings JS, Weinstein JM (2008) Information, school choice, and academic achievement: Evidence from two experiments. *Quart. J. Econom.* 123(4):1373–1414.

He Y (2014) Gaming the Boston school choice mechanism in Beijing. Working paper, Rice University, Houston, TX.

Hedblom D, Hickman BR, List JA (2019) Toward an understanding of corporate social responsibility: Theory and field experimental evidence. Working paper, NBER, Cambridge, MA.

Heyman JE, Orhun Y, Ariely D (2004) Auction fever: The effect of opponents and quasi-endowment on product valuations. *J. Interactive Marketing* 18(4):7–21.

Holmstrom B (1982) Moral hazard in teams. *Bell J. Econom.* 13(2):324–340.

Hossain T (2008) Learning by bidding. *RAND J. Econom.* 39(2):509–529.

Hossain T, List JA (2012) The behavioralist visits the factory: Increasing productivity using simple framing manipulations. *Management Sci.* 58(12):2151–2167.

Houser D, Wooders J (2005) Hard and soft closes: A field experiment on auction closing rules. Zwick R, Rapoport A, eds. *Experimental Business Research,* Vol. II (Springer, New York), 123–131.

Hoxby C, Avery C (2013) The missing "one-offs": The hidden supply of high-achieving, low-income students. *Brookings Papers Econom. Activity* 2013(1):1–65.

Hoxby C, Turner S (2015) What high-achieving low-income students know about college. *Amer. Econom. Rev.* 105(5):514–517.

Imas A, Sadoff S, Samek A (2017) Do people anticipate loss aversion? *Management Sci.* 63(5):1271–1284.

Jiang M (2016) When do stable matching mechanisms fail? The role of standardized tests in college admissions. Working paper, University of Michigan, Ann Arbor.

Jagannathan R (2019) On frequent batch auctions for stocks. Working paper, NBER, Cambridge, MA.

Kagel JH, Roth AE (2000) The dynamics of reorganization in matching markets: A laboratory experiment motivated by a natural experiment. *Quart. J. Econom.* 115(1):201–235.

Kahneman D, Lovallo D (1993) Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Sci.* 39(1):17–31.

Kahneman D, Tversky A (1979) Prospect theory: An analysis of decision under risk. *Econometrica* 47(2):263–291.

Kahneman D, Knetsch JL, Thaler RH (1990) Experimental tests of the endowment effect and the Coase theorem. *J. Political Econom.* 98(6):1325–1348.

Katok E, Roth AE (2004) Auctions of homogeneous goods with increasing returns: Experimental comparison of alternative "Dutch" auctions. *Management Sci.* 50(8):1044–1063.

Kearns M, Roth A (2019) *The Ethical Algorithm* (Oxford University Press, UK).

Kerschbamer R, Neururer D, Sutter M (2016) Insurance coverage of customers induces dishonesty of sellers in markets for credence goods. *Proc. Natl. Acad. Sci. USA*. 113(27):7454–7458.

Kerschbamer R, Neururer D, Sutter M (2019) Credence goods markets and the informational value of new media: A natural field experiment. Working paper, MPI Collective Goods Discussion Paper.

Kim K, Chung K, Lim N (2019) Third-party reviews and quality provision. *Management Sci.* 65(6):2695–2716.

Kirchler M, Huber J, Stefan M, Sutter M (2016) Market design and moral behavior. *Management Sci.* 62(9):2615–2625.

Klein TJ, Lambertz C, Stahl KO (2016) Market transparency, adverse selection, and moral hazard. *J. Political Econom.* 124(6):1677–1713.

Klemperer P (2004) *Auctions: Theory and Practice* (Princeton University Press, NJ).

Krishna A, Ünver MU (2008) Improving the efficiency of course bidding at business schools: Field and laboratory studies. *Marketing Sci.* 27(2):262–282.

Ku G, Malhotra D, Murnighan JK (2005) Toward a competitive arousal model of decision-making: A study of auction fever in live and Internet auctions. *Organization Behavior Human Decision Processes* 96(2):89–103.

Kube S, Maréchal MA, Puppe C (2012) The currency of reciprocity: Gift exchange in the workplace. *Amer. Econom. Rev.* 102(4):1644–1662.

Kube S, Maréchal MA, Puppe C (2013) Do wage cuts damage work morale? Evidence from a natural field experiment. *J. Eur. Econom. Assoc.* 11(4):853–870.

Kwasnica AM, Ledyard JO, Porter D, DeMartini C (2005) A new and improved design for multi-object iterative auctions. *Management Sci.* 51(3):315–518.

Kyle AS, Lee J (2017) Toward a fully continuous exchange. *Oxford Rev. Econom. Policy* 33(4):650–675.

Lai F, Sadoulet E, de Janvry A (2009) The adverse effects of parents' school selection errors on academic achievement: Evidence from the Beijing open enrollment program. *Econom. Edu. Rev.* 28(4):485–496.

Lazear E, Rosen S (1981) Rank-order tournaments as optimum labor contracts. *J. Political Econom.* 89(5):841–864.

Ledyard JO (1993) The design of coordination mechanisms and organizational computing. *J. Organizational Comput. Electronic Commerce* 3(1):121–134.

Leider S, Roth AE (2010) Kidneys for sale: Who disapproves, and why? *Amer. J. Transplantation* 10(5):1221–1227.

Levin J (2013) The economics of Internet markets. Acemoglu D, Arellano M, Dekel E, eds. *Advances in Economics and Econometrics* (Cambridge University Press, UK), 48–75.

Levitt SD, List JA (2007) What do laboratory experiments measuring social preferences reveal about the real world? *J. Econom. Perspective* 21(2):153–174.

Levitt SD, List JA (2009) Field experiments in economics: The past, the present, and the future. *Eur. Econom. Rev.* 53(1):1–18.

Levitt SD, List JA, Neckermann S, Sadoff S (2016) The behavioralist goes to school: Leveraging behavioral economics to improve educational performance. *Amer. Econom. J. Econom. Policy* 8(4):183–219.

Li L (2010) Reputation, trust, and rebates: How online auction markets can improve their feedback mechanisms. *J. Econom. Management Strategy* 19(2):303–331.

Li LI, Xiao E (2014) Money talks: Rebate mechanisms in reputation system design. *Management Sci.* 60(8):2054–2072.

Li LI, Tadelis S, Zhou X (2016) Buying reputation as a signal of quality: Evidence from an online marketplace. Working paper, NBER, Cambridge, MA.

Lien JW, Zheng J, Zhong X (2016) Preference submission timing in school choice matching: Testing fairness and efficiency in the laboratory. *Exp. Econ.* 19(1):116–150.

List JA (2003) Does market experience eliminate market anomalies? *Quart. J. Econom.* 118(1):41–71.

List JA (2004) Neoclassical theory vs. prospect theory: Evidence from the marketplace. *Econometrica* 72(2):615–625.

List JA (2006) The behavioralist meets the market: Measuring social preferences and reputation effects in actual transactions. *J. Political Econom.* 114(1):1–37.

List JA (2011) Does market experience eliminate market anomalies? the case of exogenous market experience. *Amer. Econom. Rev.* 101(3):313–317.

List JA, Momeni F (2020) When corporate social responsibility backfires: Evidence from a natural field experiment. *Management Sci.*, ePub ahead of print June 8, https://doi.org/10.1287/mnsc.2019.3540.

List JA, Rasul I (2011) Field experiments in labor economics. Ashenfelter O, Card D, eds. *Handbook of Labor Economics* (Elsevier, New York), 103–228.

List J, van Soest D, Stoop J, Zhou H (2020) On the role of group size in tournaments: Theory and evidence from laboratory and field experiments. *Management Sci.*, ePub ahead of print March 12, https://doi.org/10.1287/mnsc.2019.3441.

Liu TX, Yang J, Adamic L, Chen Y (2014) Crowdsourcing with all-pay auctions: A field experiment on Taskcn. *Management Sci.* 60(8):2020–2037.

Luca M, Zervas G (2016) Fake it till you make it: Reputation, competition, and Yelp review fraud. *Management Sci.* 62(12): 3412–3427.

Malaga R, Porter D, Ord K, Montano B (2010) A new end-of-auction model for curbing sniping. *J. Oper. Res. Soc.* 61(8):1265–1272.

Masterov DV, Mayer UF, Tadelis S (2015) Canary in the e-commerce coal mine: Detecting and predicting poor experiences using buyer-to-seller messages. *Proc. 16th ACM Conf. on Economics and Computation* (ACM, New York), 81–93.

McAfee RP, McMillan J (1996) Analyzing the airwaves auction. *J. Econom. Perspective* 10(1):159–175.

Menkveld A, Zoican M (2014) Need for speed? Exchange latency and market quality. Working paper, VU University, Amsterdam.

Milgrom PR (2004) *Putting Auction Theory to Work* (Cambridge University Press, UK).

Milgrom PR (2017) *Discovering Prices: Auction Design in Markets with Complex Constraints* (Columbia University Press, New York).

Milgrom PR (2019) Auction market design: Recent innovations. *Annu. Rev. Econom.* 11:383–405.

Milgrom PR, Tadelis S (2018) How artificial intelligence and machine learning can impact market design. Agrawal A, Gans J, Goldfarb A, eds. *The Economics of Artificial Intelligence: An Agenda*. National Bureau of Economic Research conference report. (University of Chicago Press), 567–585.

Mongell S, Roth AE (1991) Sorority rush as a two-sided matching mechanism. *Amer. Econom. Rev.* 81(3):441–464.

Morgan J, Orzen H, Sefton M (2012) Endogenous entry in contests. *Econom. Theory*. 51(2):435–463.

Morris S (2019) Coordination and incomplete information. *2019 Eur. Meeting of the Econometric Society Presidential Address*. MIT Working Paper.

Mostagir M, Chen Y, Yeckehzaare Y (2019). Information provision in dynamic contests: An experimental study. Working paper, University of Michigan, Ann Arbor.

Müller W, Schotter A (2010) Workaholics and dropouts in organizations. *J. Eur. Econom. Assoc.* 8(4):717–743.

Mussweiler T, Ockenfels A (2013) Similarity increases altruistic punishment in humans. *Proc. Natl. Acad. Sci. USA* 110(48):19318–19323.

Neuhoff K, Ritter N, Salah-Abou-El-Enien A, Vassilopoulos P (2016) Intraday markets for power: Discretizing the continuous trading? Working paper, SSRN.

Nie H (2007) Is a high score enough? An analysis of strategies under the college admissions mechanism. *Southern Econom.* 7:23–36.

Nikiforakis N (2008) Punishment and counter-punishment in public good games: Can we really govern ourselves? *J. Public Econom.* 92(1–2):91–112.

Nobel Foundation (2019a) The Sveriges Riksbank Prize in Economic Sciences in memory of Alfred Nobel 1994 (Nobel Media AB). Accessed May 20, 2020, https://www.nobelprize.org/prizes/economic-sciences/1994/summary/.

Nobel Foundation (2019b) The Sveriges Riksbank Prize in Economic Sciences in memory of Alfred Nobel 2002 (Nobel Media AB). Accessed May 20, 2020, https://www.nobelprize.org/prizes/economic-sciences/2002/summary/.

Nobel Foundation (2019c). The Sveriges Riksbank Prize in Economic Sciences in memory of Alfred Nobel 2012 (Nobel Media AB). Accessed May 20, 2020, https://www.nobelprize.org/prizes/economic-sciences/2012/summary/.

Nobel Foundation (2019d) The Sveriges Riksbank Prize in Economic Sciences in memory of Alfred Nobel 2017 (Nobel Media AB). Accessed May 20, 2020, https://www.nobelprize.org/prizes/economic-sciences/2017/summary/.

Nosko C, Tadelis S (2015) The limits of reputation in platform markets: An empirical analysis and field experiment. Working paper, NBER, Cambridge, MA.

Noussair CN, Tucker S (2013) Experimental research on asset pricing. *J. Econom. Survey* 27(3):554–569.

Ockenfels A, Resnick P (2012) Negotiating reputations. Croson R, Bolton GE, eds. *The Oxford Handbook of Economic Conflict Resolution* (Oxford University Press, New York), 223–237.

Ockenfels A, Roth AE (2006) Late and multiple bidding in second price Internet auctions: Theory and evidence concerning different rules for ending an auction. *Games Econom. Behavior* 55(2):297–320.

Ockenfels A, Roth AE (2013) Ending rules in Internet auctions. Vulkan N, Roth AE, Neeman Z, eds. *Handbook of Market Design* (Oxford University Press, New York), 324–344.

Ockenfels A, Selten R (2005) Impulse balance equilibrium and feedback in first price auctions. *Games Econom. Behavior* 51(1):155–170.

Ockenfels A, Sliwka D, Werner P (2015a) Bonus payments and reference point violations. *Management. Sci.* 61(7):1496–1513.

Ockenfels A, Sliwka D, Werner P (2015b) Timing of kindness: Evidence from a field experiment. *J. Econom. Behavior Organ.* 111:79–87.

Olson M, Porter D (1994) An experimental examination into the design of decentralized methods to solve the assignment problem with and without money. *Econom. Theory* 4(1):11–40.

Orrison A, Schotter A, Weigelt K (2004) Multiperson tournaments: An experimental examination. *Management Sci.* 50(2):268–279.

Ostrom E, Walker J, Gardner R (1992) Covenants with and without a sword: Self-governance is possible. *Amer. Political Sci. Rev.* 86(2):404–417.

Parkes DC, Seuken S (2020) *Algorithmic Economics: A Design Approach* (Cambridge University Press, UK).

Pathak PA, Sönmez T (2008) Leveling the playing field: Sincere and sophisticated players in the Boston mechanism. *Amer. Econom. Rev.* 98(4):1636–1652.

Pathak PA, Sönmez T (2013) School admissions reform in Chicago and England: Comparing mechanisms by their vulnerability to manipulation. *Amer. Econom. Rev.* 103(1):80–106.

Pekec A, Rothkopf MH (2003) Combinatorial auction design. *Management Sci.* 49(11):1485–1503.

Peters M, Severinov S (2006) Internet auctions with many traders. *J. Econom. Theory* 130(1):220–245.

Rabin M (2003) The Nobel Memorial Prize for Daniel Kahneman. *Scandinavian J. Econom.* 105(2):157–180.

Rasmusen EB (2006) Strategic implications of uncertainty over one's own private value in auctions. *B.E. J. Theoret. Econom.* 6(1):223–237.

Rassenti SJ, Smith VL, Bulfin RL (1982) A combinatorial auction mechanism for airport time slot allocation. *Bell J. Econom.* 13:402–417.

Roth AE (1982) The economics of matching: Stability and incentives. *Math. Oper. Res.* 7(4):617–628.

Roth AE (1990) New physicians: A natural experiment in market organization. *Science* 250(4987):1524–1528.

Roth AE (1991) A natural experiment in the organization of entry-level labor markets: Regional markets for new physicians and surgeons in the United Kingdom. *Amer. Econom. Rev.* 81(3):415–440.

Roth AE (2002) The economist as engineer: Game theory, experimentation, and computation as tools for design economics. *Econometrica* 70(4):1341–1378.

Roth AE (2007) Repugnance as a constraint on markets. *J. Econom. Perspective* 21(3):37–58.

Roth AE (2008) What have we learned from market design? *Econom. J. (London)* 118(527):285–310.

Roth AE (2015) Experiments in market design. Kagel JH, Roth AE, eds. *Handbook of Experimental Economics* (Princeton University Press, NJ), 290–346.

Roth AE, Ockenfels A (2002) Last-minute bidding and the rules for ending second-price auctions: Evidence from eBay and amazon auctions on the Internet. *Amer. Econom. Rev.* 92(4):1093–1103.

Roth AE, Peranson E (1997) The effects of the change in the NRMP matching algorithm. *JAMA* 278(9):729–732.

Roth AE, Peranson E (1999) The redesign of the matching market for American physicians: Some engineering aspects of economic design. *Amer. Econom. Rev.* 89(4):748–779.

Roth AE, Wilson RB (2019) How market design emerged from game theory: A mutual interview. *J. Econom. Perspective* 33(3):118–143.

Roth AE, Xing X (1994) Jumping the gun: Imperfections and institutions related to the timing of market transactions. *Amer. Econom. Rev.* 84(4):992–1044.

Rothkopf MH, Teisberg TJ, Kahn EP (1990) Why are Vickrey auctions rare? *J. Political Econom.* 98(1):94–109.

Servaes H, Tamayo A (2013) The impact of corporate social responsibility on firm value: The role of customer awareness. *Management Sci.* 59(5):1045–1061.

Sheremeta RM (2010) Experimental comparison of multi-stage and one-stage contests. *Games Econom. Behavior* 68(2):731–747.

Sheremeta RM (2011) Contest design: An experimental investigation. *Econom. Inquiry* 49(2):573–90.

Sliwka D, Werner P (2017) Wage increases and the dynamics of reciprocity. *J. Labor Econom.* 35(2):299–344.

Smith VL (1982) Microeconomic systems as an experimental science. *Amer. Econom. Rev.* 72(5):923–955.

Sönmez T, Ünver MU (2010) Course bidding at business schools. *Internat. Econom. Rev.* 51(1):99–123.

Tadelis S (2016) Reputation and feedback systems in online platform markets. *Annu. Rev. Econom.* 8(1):321–340.

Takeuchi K, Lin JC, Chen Y, Finholt T (2010) Scheduling with package auctions. *Experiment. Econom.* 13:476–499.

Thaler RH, Johnson EJ (1990) Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management Sci.* 36(6):643–660.

Tonin M, Vlassopoulos M (2015) Corporate philanthropy and productivity: Evidence from an online real effort experiment. *Management Sci.* 61(8):1795–1811.

Varian H (2007) Position auctions. *Internat. J. Industrial Organization* 25(6):1163–1178.

Varian H (2009) Online ad auctions. *Amer. Econom. Rev.* 99(2):430–434.

Vickrey WS (1961) Counterspeculation, auctions and competitive sealed tenders. *J. Finance* 16:8–37.

Waddock SA, Graves SB (1997) The corporate social performance-financial performance link. *Strategic Management J.* 18(4):303–319.

Wei L (2015) Chinese college admissions reform: Natural experiments of matching. Working paper, Wuhan University, China.

Wilcox RT (2000) Experts and amateurs: The role of experience in Internet auctions. *Marketing Lett.* 11(4):363–374.

Wilson R (1985) Reputations in games and markets. Roth AE, ed. *Game-Theoretic Models of Bargaining* (Cambridge University Press, New York), 65–84.

Wilson R (2002) Architecture of power markets. *Econometrica* 70(4):1299–1340.

Wolf JR Jr, Arkes HR, Muhanna WA (2005) Is overbidding in online auctions the result of a pseudo-endowment effect? Working paper, Ohio State University, Columbus.

Wu B, Zhong X (2014) Matching mechanisms and matching quality: Evidence from a top university in China. *Games Econom. Behavior* 84:196–215.

Zhang S, Riordan R (2011) Technology and market quality: The case of high frequency trading. *ECIS 2011 Proc, Helsinki, Finland, January 7, 2011* (Association for Information Systems, Atlanta).